张景中／主编

走进教育数学

Go to Educational Mathematics

微积分快餐

林　群／著

科学出版社
www.sciencep.com

"十一五"国家重点图书出版规划项目

走进教育数学

Go to Educational Mathematics

微积分快餐

科学出版社

北京

内 容 简 介

本书寻找最少且自封（不依赖于未证明的结果）的微积分，即最少的概念：微分和积分（实是一个概念，后者乃前者之和）；最少的定理：基本定理和泰勒定理（实是一个定理，后者乃前者的连用）；最简的解释（实是两张图）、最短的证明（实是两行算术，没有更多）、最少的数学符号（阿基米德的传统，多用文字和图形）．这些概念、定理和证明只用到两张图、两行算术，不用实数，适合于文科；对理科还要加上最少的（即一个）微分方程，这时才用到实数．

简言之，最少的微积分 = 两个（或一个）概念 + 两个（或一个）定理 + 一个方程．归根结底，就是两张图、两行算术，加上一点实数，没有更多．

图书在版编目(CIP)数据

微积分快餐／林群著．—北京：科学出版社，2009
（走进教育数学／张景中主编）
ISBN 978-7-03-025043-8

Ⅰ．微…　Ⅱ．林…　Ⅲ．微积分－教学研究　Ⅳ．O172-4

中国版本图书馆 CIP 数据核字（2009）第 120428 号

丛书策划：李　敏
责任编辑：李　敏／责任校对：钟　洋
责任印制：钱玉芬／整体设计：黄华斌

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

北京佳信达欣艺术印刷有限公司 印刷
科学出版社发行　各地新华书店经销
＊
2009 年 8 月第 一 版　开本：B5（720×1000）
2009 年 8 月第一次印刷　印张：10 1/4　插页：2
印数：1—6 000　字数：100 000

定价：25.00 元

总 序

看到本丛书，多数人会问这样的问题：

“什么是教育数学?”

“教育数学和数学教育有何不同?”

简单说，改造数学使之更适宜于教学和学习，是教育数学为自己提出的任务.

把学数学比作吃核桃．核桃仁美味而富有营养，但要砸开才能吃到它．有些核桃，外壳与核桃仁紧密相依，成都人形象地叫它们“夹米子核桃”，如若砸不得法，砸开了还很难吃到．数学教育要研究的，就是如何砸核桃吃核桃．教育数学呢，则要研究改良核桃的品种，让核桃更美味，更营养，更容易砸开吃净.

“教育数学”的提法，最早出现在笔者1989年所写的《从数学教育到教育数学》中．其实，教育数学的活动早已有之，如欧几里得著《几何原本》，柯西写《分析教程》，

都是教育数学的经典之作.

数学教育有很多世界公认的难点，如初等数学里的几何和三角，高等数学里面的微积分，都比较难学. 为了对付这些难点，很多数学老师、数学教育专家前赴后继，做了大量的研究，写了很多的著作，进行了广泛的教学实践. 多年实践，几番改革，还是觉得太难，不得不“忍痛割爱”，少学或者不学. 教育数学则从另一个角度看问题：这些难点的产生，是不是因为前人留下来的知识组织得不够好，不适于数学的教与学？能不能优化数学，改良数学，让数学知识变得更容易学习呢？

知识的组织方式和学习的难易有密切的联系. 英语中12个月的名字：January，February，……. 背单词要花点工夫吧？如果改良一下：一月就叫 Monthone，二月就叫 Monthtwo，等等，马上就能理解，就能记住，学起来就容易多了. 生活的语言如此，科学的语言——数学——何尝不是这样呢？

很多人认为，现在小学、中学到大学里所学的数学，从算术、几何、代数、三角到微积分，都是几百年前甚至几千年前创造出来的数学. 这些数学的最基本的部分，普遍认为是经过千锤百炼，相当成熟了. 对于这样的数学内容，除了选择取舍，除了教学法的加工之外，还有优化改革的余地吗？

但事情还可以换个角度看. 这些进入了课堂的数学，是在不同的年代，不同的地方，由不同的人，为不同的目的而创造出来的，而且其中很多不是为了教学的目的而创造出来的. 难道它们会自然而然地配合默契，适宜于教学和学

习吗?

看来，这主要不是一个理论问题，而是一个实践问题.

走进教育数学，看看教育数学在做什么，有助于回答这类问题.

随便翻翻这几本书，就能了解教育数学领域里近20年来做了哪些工作.从已有的结果看到，教育数学有事可做，而且能做更多的事情.

比如微积分教学的改革，这是在世界范围内被广为关注的事.丛书中有两本专讲微积分，主要还不是讲教学方法，而是讲改革微积分本身.

由牛顿和莱布尼茨创建的微积分，是第一代的微积分.这是说不清楚的微积分.创建者说不清楚，使用微积分解决问题的数学家也说不清楚.原理虽然说不清楚，应用仍然在蓬勃发展.微积分在说不清楚的情形下发展了130多年.

柯西和魏尔斯特拉斯等建立了严谨的极限理论，巩固了微积分的基础，形成了第二代的微积分.数学家把微积分说清楚了.但是由于概念和推理繁琐迂回，对于绝大多数学习高等数学的人来说，是听不明白的微积分.微积分在多数学习者听不明白的情形下，又发展了170多年，直到今天.

第三代的微积分，是正在创建发展的新一代的微积分.人们希望微积分不但严谨，而且直观易懂，简易明快.让学习者用较少的时间和精力就能够明白其原理，不但知其然而且知其所以然.不但数学家说得清楚，而且非数学专业的多数学子也能听得明白.

第一代微积分和第二代微积分，在具体计算方法上基本相同；不同的是对原理的说明，前者说不清楚，后者说

清楚了.

第三代微积分和前两代微积分，在具体计算方法上也没有不同；不同的仍是对原理的说明.

几十年来，国内外都有人从事第三代微积分的研究以至教学实践. 这方面的努力，已经有了显著的成效. 在我国，林群院士近10年来在此方向做了大量的工作. 本丛书中的《微积分快餐》，就是他在此领域的代表作.

古今中外，通俗地介绍微积分的读物极多，但能够兼顾严谨与浅显直观的几乎没有.《微积分快餐》做到了. 一张图，一个不等式，几行文字，浓缩了微积分的精华. 作者将微积分讲得轻松活泼、简单明了而且严谨自封，让读者在品尝快餐的过程中进入了高等数学的殿堂.

丛书中还有一本《直来直去的微积分》，是笔者学习微积分的心得. 书中从"瞬时速度有时比平均速度大，有时比平均速度小"这个平凡的陈述出发，不用极限和实数，"微分不微，积分不积"，直截了当地建立了微积分基础理论. 书中概念与《微积分快餐》中的逻辑等价而呈现形式不尽相同，殊途同归，显示出第三代微积分的丰富多彩.

回顾历史，牛顿和拉格朗日都曾撰写著作，致力于建立不用极限也不用无穷小的微积分，或证明微积分的方法，但没有成功. 我国数学大师华罗庚所撰写的《高等数学引论》中，也曾刻意求新，不用中值定理或实数理论而寻求直接证明"导数正则函数增"这个具有广泛应用的微积分基本命题，可惜也没有达到目的.

前辈泰斗是我们的先驱. 教育数学的进展实现了先驱们简化微积分理论的愿望.

两本关于微积分的书，都专注于基本思想和基本概念的变革. 基本思想、基本概念，以及在此基础上建立的基本定理和公式，是这门数学的筋骨. 数学不能只有筋骨，还要有血有肉. 中国高等教育学会教育数学专业委员会理事长、全国名师李尚志教授的最新力作《数学的神韵》，是有血有肉、丰满生动的教育数学. 书中的大量精彩实例可能是你我熟悉的老故事，而作者却能推陈出新，用新的视角和方法处理老问题，找出事物之间的联系，发现不同中的相同，揭示隐藏的规律. 幽默的场景，诙谐的语言，使人在轻松阅读中领略神韵，识破玄机. 看看这些标题，“简单见神韵”、“无招胜有招”、“茅台换矿泉”、“凌波微步微积分”，可以想见作者的功力非同一般！特别值得一提的是，书中对微积分的精辟见解，如用代数观点演绎无穷小等，适用于第一代、第二代和第三代微积分的教学与学习，望读者留意体味.

练武功的上乘境界是“无招胜有招”，但武功仍要从一招一式入门. 解数学题也是如此. 著名数学家和数学教育家项武义先生说，教数学要教给学生“大巧”，要教学生“运用之妙，存乎一心”，以不变应万变，不讲或少讲只能对付一个或几个题目的“小巧”. 我想所谓“无招胜有招”的境界，就是“大巧”吧！但是，小巧固不足取，大巧也确实太难. 对于大多数学子，还要重视有章可循的招式，由小到大，以小御大，小题做大，小中见大. 朱华伟教授和钱展望教授的《数学解题策略》，踏踏实实地从一招一式一题一法着手，探秘发微，系统地阐述数学解题法门，是引领读者登堂入室之作. 作者是数学奥林匹克领域的专家. 数学奥林

匹克讲究题目出新，不落老套. 我看了这本书里的不少例题，看不出有哪些似曾相识，真不知道他是从哪里搜罗来的！

朱华伟教授还为本丛书写了一本《从数学竞赛到竞赛数学》. 竞赛数学当然就是奥林匹克数学. 华伟教授认为，竞赛数学是教育数学的一部分. 这个看法是言之成理的. 数学要解题，要发现问题、创造方法. 年复一年进行的数学竞赛活动，不断地为数学问题的宝库注入新鲜血液，常常把学术形态的数学成果转化为可能用于教学的形态. 早期的国际数学奥林匹克试题，有不少进入了数学教材，成为例题和习题. 竞赛数学与教育数学的关系，于此可见一斑.

写到这里，忍不住要为数学竞赛说几句话. 有一阵子，媒体上出现不少讨伐数学竞赛的声音，有的教育专家甚至认为数学竞赛之害甚于黄赌毒. 我看了有关报道后第一个想法是，中国现在值得反对的事情不少，论轻重缓急还远远轮不到反对数学竞赛吧. 再仔细读这些反对数学竞赛的意见，可以看出来，他们反对的实际上是某些为牟利而又误人子弟的数学竞赛培训. 就数学竞赛本身而言，是面向青少年中很小一部分数学爱好者而组织的活动. 这些热心参与数学竞赛的数学爱好者（还有不少数学爱好者参与其他活动，例如青少年创新发明活动、数学建模活动、近年来设立的丘成桐中学数学奖），估计不超过约两亿中小学生的百分之五. 从一方面讲，数学竞赛培训活动过热产生的消极影响，和升学考试体制以及教育资源分配过分集中等多种因素有关，这笔账不能算在数学竞赛头上；从另一方面看，大学招生和数学竞赛挂钩，也正说明了数学竞赛活动的成功因

而得到认可. 对于青少年的课外兴趣活动, 积极的对策不应当是限制堵塞, 而是开源分流. 发展多种课外活动, 让更多的青少年各得其所, 把各种活动都办得像数学竞赛这样成功并且被认可, 数学竞赛培训活动过热的问题自然就化解或缓解了.

回到前面的话题. 上面说到"大巧"和"小巧", 自然想到还有"中巧". 大巧法无定法, 小巧一题一法. 中巧呢, 则希望用一个方法解出一类题目. 也就是说, 把数学问题分门别类, 一类一类地寻求可以机械执行的方法, 即算法. 中国古代的《九章算术》, 就贯穿了分类解题寻求算法的思想. 中小学里学习四则算术、代数方程, 大学里学习求导数, 学的多是机械的算法. 但是, 自古以来几何命题的证明却千变万化, 法无定法. 为了找寻几何证题的一般规律, 从欧几里得、笛卡儿到希尔伯特, 前赴后继, 孜孜以求. 我国最高科技奖获得者、著名数学家吴文俊院士指出, 希尔伯特是第一个发现了几何证明机械化算法的人. 在《几何基础》这部名著中, 希尔伯特对于只涉及关联性质的这类几何命题, 给出了机械化的判定算法. 由于受时代的局限性, 希尔伯特这一学术成果并不为太多人所知. 直到 1977 年, 吴文俊先生提出了一个新的方法, 可以机械地判定初等几何中等式型命题的真假. 这一成果在国际上被称为"吴方法", 它在几何定理机器证明领域中掀起了一个高潮, 使这个自动推理中最不成功的部分变成了最成功的部分.

吴方法和后来提出的多种几何定理机器证明的算法, 都不能给出人们易于检验和理解的证明, 即所谓可读证明. 国内外的专家一度认为, 机器证明的本质在于"用量的复杂

克服质的困难"，所以不可能机械地产生可读证明.

笔者基于1974年在新疆教初中时指导学生解决几何问题的心得，总结出用面积关系解题的规律. 在这些规律的基础上，1992年提出消点算法，和周咸青、高小山两位教授合作，创建了可构造等式型几何定理可读证明自动生成的理论和方法，并在计算机上实现. 最近在网上看到，面积消点法也多次在国外的不同的系统中实现了. 本丛书中的《几何新方法和新体系》，包括了面积消点法的通俗阐述，以及笔者提出的一个有关面积方法的公理系统，由冷拓同志协助笔者整理成书. 教育数学研究的副产品解决了机器证明领域中的难题，对笔者而言实属侥幸.

基于对数学教育的兴趣，笔者从1974年以来，在30多年间持续地探讨面积解题的规律，想把几何变容易一些. 后来发现，国内外的中学数学教材里，已经把几何证明删得差不多了. 于是"迷途知返"，把三角作为研究的重点. 数学教材无论如何改革，三角总是删不掉的吧. 本丛书中的《一线串通的初等数学》，讲的是如何在小学数学知识的基础上建立三角，以三角的发展引出代数工具并探索几何，把三者串在一起的思路.

在《一线串通的初等数学》中没有提到向量. 其实，向量早已下放到中学，与传统的初等数学为伍了. 在上海的数学教材里甚至在初中就开始讲向量. 讲了向量，自然想试试用向量解决几何问题，看看向量解题有没有优越性. 可惜在教材里和刊物上出现的许多向量例题中，方法略嫌繁琐，反而不如传统的几何方法简捷优美. 如何用向量法解几何题? 能不能在大量的几何问题的解决过程中体现向量解题

的优越性？这自然是教育数学应当关心的一个问题. 为此，本丛书推出一本《绕来绕去的向量法》. 书中用大量实例说明，如果掌握了向量解题的要领，在许多情形下，向量法比纯几何方法或者坐标法干得更漂亮. 这要领，除了向量的基本性质，关键就是"回路法". 绕来绕去，就是回路之意. 回路法是笔者的经验谈，没有考证前人是否已有过，更没有上升为算法. 书稿主要由彭翕成同志执笔，绝大多数例子，也是他采集加工的.

谈起中国的数学科普，谈祥柏的名字几乎无人不知. 老先生年近八旬，从事数学科普创作超过半个世纪，出书 50 多种，文章逾千篇. 对于数学的执著和一生的爱，洋溢于他为本丛书所写的《数学不了情》的字里行间. 哪怕仅仅信手翻上几页，哪怕是对数学知之不多的中小学生，也会被一个个精彩算例所显示的数学之美和数学之奇深深吸引. 书中涉及的数学知识似乎不多不深，所蕴含的哲理却足以使读者掩卷遐想. 例如，书中揭示出高等代数的对称、均衡与和谐，展现了古老学科的青春；书中提到海峡两岸的数学爱好者发现了千百年来从无数学者、名人的眼皮底下滑过去的"自然数高次方的不变特性"，这些生动活泼的素材，兼有冰冷的思考与火热的激情，无论读者偏文偏理，均会有所收益.

沈文选教授长期从事中学数学研究、初等数学研究、奥林匹克数学研究和教育数学的研究. 他的《走进教育数学》和本丛书同名，是一本从学术理论角度探索教育数学的著作. 在书中他试图诠释"教育数学"的概念，探究"教育数学"的思想源头与内涵；提出"整合创新优化"、

"返璞归真优化"等优化数学的方法和手段；并提供了丰富的案例. 笔者原来杜撰出"教育数学"的概念，虽然有些实例，但却凌乱无序，不成系统. 经过文选教授的旁征博引，诠释论证，居然有了粗具规模的体系框架，有点学科模样了. 这确是意外的收获.

浏览着这风格不同并且内容迥异的10本书，教育数学领域的现状历历在目. 这是一个开放求新的园地，一个蓬勃发展的领域. 在这里耕耘劳作的人们，想的是教育，做的是数学，为教育而研究数学，通过丰富发展数学而推进教育. 在这里大家都做自己想做的事，提出新定义新概念，建立新方法新体系，发掘新问题新技巧，寻求新思路新趣味，凡此种种，无不是为教育而做数学.

为教育而做数学，做出了些结果，出了这套书，这仅仅是开始. 真正重要的是进入教材，进入课堂，产生实效，让千千万万学子受益，进而推动社会发展，造福人类. 这才是作者们和出版者的大期望. 切望海内外同道者和不同道者指正批评，相与切磋，共求真知，为数学教育的进步贡献力量.

张景中

2009年7月

前 言

数学需要创造性的读者，每段都需要严格的拷问和推敲.

我教过微积分，觉得太多、太难了. 内容即使能够背下，也是知其然而不知其所以然①，只能照本宣科，或拿结论做题. 如果教师自己一知半解，学生岂不更加迷糊？所以我一直有改变其状态的愿望. 直到有一天，我在一棵古树下散步，听到导游对观光团讲解：这棵古树年年在长高，年年有测量员来测高……本是日常事，却是说者无心听者有意：树高是怎么测的？必须砍树吗？学过中学三角就知道，有了斜率的概念，无需砍树(或爬树)也能测

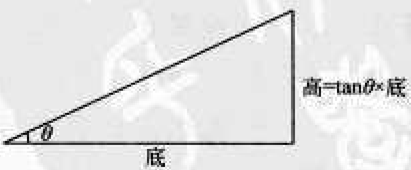


图 0-1 中学直边三角形（酵母）

① 其他教师也有同感，这说明现有教材把微积分写得太难了.

出树高(图 0-1). 多轻松，谁说斜率(数学概念)没有用！更有用的是，由此拾到微积分的酵母(或灵感)，认出它的真面目：只要把直角三角形的斜边弯起来，改成在山坡上测山高(图 0-2). 便可由中学三角(直边)进入大学微积分(曲边)甚至微分方程①.

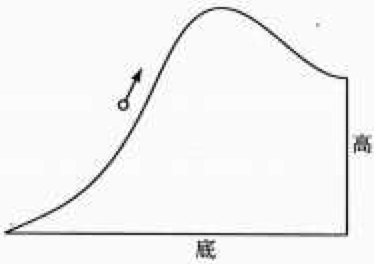


图 0-2　大学曲边三角形（发酵）已知函数（山坡）的导函数（斜率）求函数自身（山高）.

这不就是用斜率求山高吗？这不就是曲边三角吗？所以，大学和中学既不同(曲和直)又统一(都是三角学，都用斜率求高)！

怎么想得到，一位导游竟成了我的微积分启蒙老师：教室一学期，树下一分钟！② 圣人(牛顿)曾在苹果树下发现了万有引力，凡人在古树下解读了微积分. 树下这个心得，微积分不过是求高，值得推广. 我曾到一些院校交流过，并在一位朋友的激励下，把心得发表在光明日报和人民日报上(1997 年). 总之，照着一张求高图一眼认出微积分的真面目. 图形在先，证明在后③，没有想到，一个小诀窍，能用两行算术一举证明了基本定理(长了记不住)，泰勒定

① 这是最简单的一个微分方程. 还有最实际的一个微分方程，见本书 1.12 节，这时要用实数（大学数学）.

② 我在教室学的是求面积（先变成原函数再求导，这是间接法，凡人想不到，天上掉下来的），在树下学的是求高（所以利用斜率，这是直接法，凡人也想得到，是必然的结果）.

③ 数学家吴文俊说得好，直觉和推理都需要，但是只会推理，缺乏直觉，是不会有创造性的.

理的真面目也一眼认出(不过是基本定理的连用)，无需另找证明. 微积分的核心概念(微分和积分)和关键技术(基本定理和泰勒定理)，由一学期、几百页压到几学时、几页，使物理和数学同步教学.

本书只有图形和算术，无需更多(如实数,或那些难以证明的结果)，可能是最小的代价. 怎么会有这么轻松的事? 可信吗? 严格吗? 准确吗? 是不是胡闹? 好在数学中的真伪是能判断的. 当读者仔细研读了本书头几页(只要头几页)后，不必相信一面之词，自己能做出真伪的判断.

本书突出重点，不计其余，没有太多太难的东西(除了第5章微分方程，都不用实数)，只是必需品，不奢侈 (功能少，短平快).

林群

2008年4月

目　录

第 1 章
白话微积分(轻松但严格)

世上只有数学学科才是严格的，太难得，丢掉严格等于丢掉人类最宝贵的财富！微积分达到严格和轻松的平衡.

大学和中学的数学差别在于曲和直. 怎样阐明这差别(最后消除差别以直解读曲)？本章前两节的几页中，通过白话和图形，加上几行算术，来完成这任务，多轻松！

由于白话把数学人文化了，放弃形式推理，又兼顾严格性，显得不三不四，初读难知所云，更需要创造性的思考和拷问. 但是，如果通过思考能读几页就入门，值得！

1.1　切线和微分:计算相对误差

先有图形：把曲线小段用切线来代替，出现误差；再有严格

处理：计算更有用的“相对误差”（关键词）而不是误差自身.

先有感觉：什么是曲线的切线呢？几何感觉就是在一点附近最“贴近”于曲线的那条直线(最好的逼近). 它怎么构造呢？按现有教材，为了在点 t 处决定切线，在曲线上取与 t 不同的点 s，引割线 ts，然后让 s 沿着曲线去接近 t，如果割线 ts 趋向一个极限位置，则这个极限位置的直线就叫做曲线在点 t 处的切线(图 1-1).

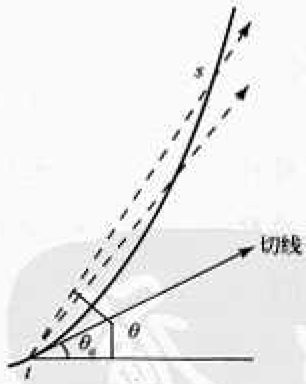


图 1-1 曲线在点 t 处的切线

何谓极限位置？这只是感觉！割线的极限位置，或切线，不一定存在. 如果它存在，就称曲线在这一点处“可微”. 本书只考虑可微的情形，即假设切线存在！

为什么切线在点 t 附近最贴近于曲线呢？让我们来计算它们之间的“距离”（图 1-2）：起点附近

$$\text{测量误差} = \text{小段高} - \text{切线高}, \tag{1-1}$$

$$\text{切线高} = \text{斜率} \times \text{底}. \tag{1-2}$$

由图 1-2 看得见，无需花脑筋计算，这个测量误差会随着曲

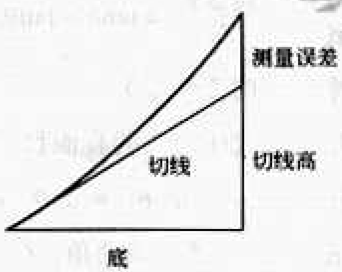


图 1-2　曲线小段求高(微分)

小段高≈切线高

线小段变小而变小. 但是, 更有用的是"相对误差", 或测量误差和所取曲线小段(或底)之比,

$$相对误差=\frac{测量误差}{底}. \tag{1-3}$$

它又是什么呢? 看得见吗? 由切线的几何构造(图1-3): 起点附近

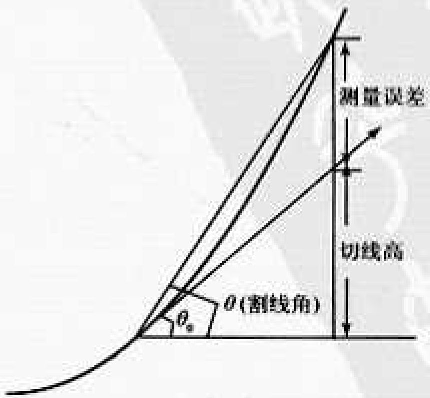


图 1-3　计算相对误差

$$\frac{\text{测量误差}}{\text{底}}=\frac{\text{小段高}}{\text{底}}-\frac{\text{切线高}}{\text{底}}=\tan\theta-\tan\theta_0=\text{割线斜率}-\text{切线斜率}$$（或平均变化率 - 瞬时变化率）.

更清楚些，通过三角关系式认出它的真面目：

$$\text{相对误差}=\left|\tan\theta-\tan\theta_0\right|\leqslant C\left|\theta-\theta_0\right|\ll 1. \tag{1-4}$$

右端已是可见的变量(其中 θ 是割线动角，C 是常数). 看得见它随着割线趋向切线(或底变小)会变得任意小(指绝对值小于任一指定的数，今后简记为 $\ll 1$). 这里我们有意用可见的量躲开正统教材中抽象的 $\varepsilon-\delta$ 语言.

相对误差(1-4)很小，说明测量误差(1-1)不是一般的小，而是非常小.

N. Arnold(1995)的动画(图 1-4)也说明了这一点：如果底很小，测量误差便非常小，这就是为什么相对误差会很小的原因.

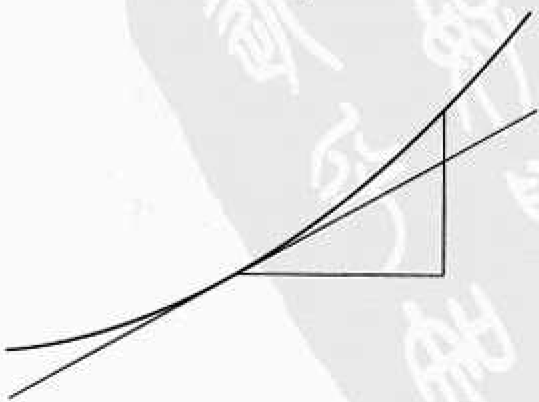


图 1-4 切线性质

测量误差与所取曲线小段相比是很小的.

先有图形后再严格：前面(图1-1)对于切线的图形感觉，只是引导，不能算数．以下反之，我们“过河拆桥”，抛开图形，利用代数公式(1-3)或(1-4)的

$$\text{相对误差} \ll 1 \tag{1-5}$$

(起点附近，下同)定义切线：切线被它惟一确定下来，(见注2)．这就是切线存在(或曲线可微)的代数定义，也是本书惟一用到的定义，缺它不行！初学者最难接受“相对误差”这个概念，总觉得别扭，但是法国数学课程标准(2001)高三部分(见(1-19))都能这么讲，我们也要有信心！

为什么要讲相对误差呢？通过它，才能划清切线和割线的界线．如果用其他直线来代替曲线，不能使起点附近相对误差变小(见注3)．只有切线，才能使它变小，所以是最好的代替．因此，微分学不是一般的以直代曲，而是选切线——这一点非常明智！

注1(初学略去)　对于(1-4)中抽象符号“$\ll 1$”还有一个解读：在常见情形(曲线不仅可微，而且更光滑，见1.11节注6)这个比值(或$|\theta-\theta_0|$)只跟曲线小段或底的大小成比例，因而随着底一起变小，底有多小它就能多小．于是，$\ll 1$就像$1/n\to 0$那样，会变得任意小，无需$\varepsilon-\delta$语言．

注2(切线性质)　(1-1)，(1-2)，(1-3)，(1-5)已含惟一性．反证法：如果曲线小段在起点有两条切线，有两个斜率A和B，得到附近不同的相对误差C和D：

$$\frac{\text{小段高}}{\text{底}}=\text{斜率}+\frac{\text{测量误差}}{\text{底}}$$

$$= A + \text{相对误差 } C$$

$$= B + \text{相对误差 } D.$$

令

$$d := A - B = \text{相对误差 } D - \text{相对误差 } C.$$

右端相对误差在起点附近能够小于定数 d，结果 $d < d$，矛盾.

注 3（初学略去） 切线的测量误差跟底相比趋于零，但是其他直线的测量误差跟底相比趋于非零常数. 看图自明（图1-5）.

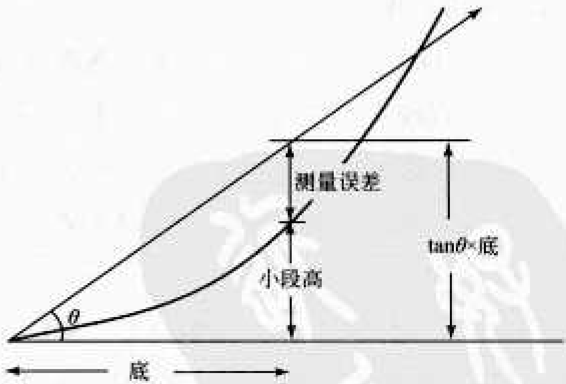


图 1-5 割线的相对误差

其他直线的斜率如果为 $\tan\theta$，则有

$$\frac{\text{测量误差}}{\text{底}} = \frac{\tan\theta\times\text{底}}{\text{底}} - \frac{\text{小段高}}{\text{底}} \to \tan\theta - \tan\theta_0 = \text{常数}, \quad \text{当底} \to 0 \quad (\theta \neq \theta_0).$$

所以，切线，也只有切线的测量误差才是最小（最好的逼近），这就是为什么把曲线小段换成切线（最好的代替）. 曲线在一点有惟一的切线密切贴近它，这就是微分学的秘诀！

以上是曲线在一点处的切线. 如果曲线在不同的点处都存在

切线，就称这曲线点点“可微”或处处“光滑”．以下(1.2节)只考虑点点可微的情形．所以曲线在各点附近的相对误差≪1．但是要注意，当曲线从一点移到另一点时，相对误差(或测量误差)变小的速度有快有慢，不见得一致．或者说，要使各点附近相对误差小于同一个指定的数，用的底有长有短，不见得一致．

相对误差的概念如此重要，需要读者创造性的思考和严格的拷问．

暂停．下面由切线转到微分．

有了切线或斜率的定义，就有了微分(或切线高)的定义：

$$微分 = 斜率 \times 底, \tag{1-6}$$

于是曲线小段的高便用微分来代替，它们之间的差量即起点附近的测量误差(1-1)．后者非常小，以至于和所取曲线小段(或底)相比，相对误差(1-4)也很小．法国数学课程标准(2001)高三部分就是这么刻画微分的(见(1-19))，必须接受它！

至此，微积分的首要任务——图1-2：曲线小段求高(微分)，计算相对误差，已经完成！

本节定义的微分是最好的逼近，它保证相对误差≪1，这为下面的积分学做好了准备．

1.2　积分学:计算相对误差的平均

1.1节考虑的是曲线小段，本节转向曲线全段，它用分段切线来代替(也叫包络)．自然地，全段(曲边三角形的)高便用分

段微分(直边三角形的高)的和来代替(图 1-6).

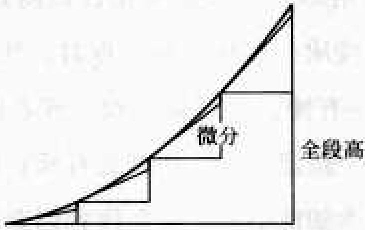


图 1-6 曲线全段求高

通常积分用来求面积，本书用来求高.

让我们计算这差量:

全段误差 = 全段高 − 微分的和 =(小段高 − 微分) 的和

=(测量误差) 的和. (1-7)

(用到(1-2)、(1-6)). 右边的和又怎么计算呢? 当分割加密时，项数会越来越多，会不会造成误差积累或扩大呢? 还是像我们所期望的，会变得任意小呢? 凭感觉不够了，有待进一步计算.

很幸运，通过两行算术，右边的和立刻改写为 1.1 节有过的“局部相对误差”的平均:

$$\frac{(\text{测量误差})\text{的和}}{\text{底的和}} := \frac{a_1 + a_2 + \cdots}{b_1 + b_2 + \cdots} \quad (\text{再拆开：一行算术})$$

$$= \left(\frac{b_1}{b_1 + b_2 + \cdots}\right)\left(\frac{a_1}{b_1}\right) + \left(\frac{b_2}{b_1 + b_2 + \cdots}\right)\left(\frac{a_2}{b_2}\right) + \cdots$$

(其中系数 >0，相加 =1，又一行算术)

$$:= \left(\frac{a_1}{b_1}, \frac{a_2}{b_2}, \cdots\right)\text{的平均}$$

$$= \left(\frac{\text{测量误差}}{\text{底}}\right)\text{的平均},$$

变为(1-3)"局部相对误差"的加权平均(当各底相等,才是算术平均,后者学生易接受)不再是求和！其中测量误差(或相对误差)不问是正是负、是大或小，只关心它们平均后的大小. 于是，由于分母(底的和)是常数 C，分子

$$(测量误差)的和 = C\cdot(相对误差)的平均$$

或(1-7)中全段误差：

$$全段高 - 微分的和 = C\cdot(相对误差)的平均. \tag{1-8}$$

这算是证明还是改写？从头到尾，只用两行算术！

幸运结论　(1-8)称为转换恒等式，把(1-7)中的求和改写为平均，而不是放大！

由(1-8)左边感觉不到误差的大小，但是右边改写为 1.1 节有过的"局部相对误差"后，才认出了全段误差的真面目. 于是，转换恒等式(1-8)给出了基本定理必要且充分的条件：当且仅当

$$|平均| \ll 1 \tag{1-9}$$

(小于任一指定的数)，才有

$$|全段高 - 微分的和| = C\cdot|平均| \ll 1. \tag{1-10}$$

此即基本定理. 何以见得？由于(1-10)中的全段高是给定的常数，微分和就是一个黎曼和，(1-10)便读作

$$|常数 - 黎曼和| \ll 1,$$

或黎曼和趋于这个常数，这就称为可积((1-9)称为可积条件). 这个常数就定义为积分：

$$全段高 = 微分的积分. \tag{1-11}$$

如果对(1-11)还有疑虑，请跟下面的图1-12比照！

下面的讨论有点啰嗦，初学略去.

基本定理(1-11)乃微积分之本，是今后一切应用之平台．认清它，一本万利．由于(1-9)是基本定理的必要且充分的可积条件，它便成为认清基本定理的出发点．所以必需仔细观察：初看起来，平均介于某两个相对误差之间，既然每个都很小，那么平均是不是就很小呢？对一般的人就差不多了；对严谨的人，仔细一想，(1-9)中的平均有无限多个，对应于各个分割．当这些分割加密，各段相对误差会不同程度地变小，不见得一致小于同一个指定的数，未必保证平均变小．幸亏，存在更简单的判别法如下：

1）曲线"一致可微"．这是指相对误差不依赖于分点的位置，一致变小，读成

$$|(\text{相对误差})\text{的上界}| \ll 1 \qquad (1\text{-}12)$$

（当底变小）．由于

$$|\text{平均}| \leqslant |(\text{相对误差})\text{在某一分点}|$$
$$\leqslant |(\text{相对误差})\text{的上界}|,$$

于是全段误差有估计：对于任何分割

$$|\text{全段高}-\text{微分的和}| \leqslant C \cdot |(\text{相对误差})\text{的上界}|. \qquad (1\text{-}13)$$

条件(1-12)正是拉克斯书中(1980)所用的条件，并不稀奇．它还等价于

2）曲线"连续可微"．

由于等价性的证明有赖于实数，所以初读先承认．

但是，条件(1-12)只是充分并非必要. 当曲线可微但不再“连续”，这些地方的相对误差就不会一致变小，但是它们的平均可能变小. 例如，

3) 曲线“可微”，但有“有限个间断点”. 这时有限个相对误差只是有界(见(3-11))，平均也能变小，于是全段误差也有估计(1-13)，但要加小扰动(见(3-13)).

4) 曲线“可微且单调”，允许有“无限个间断点”. 这时相对误差的和能够保持有界($\leqslant C$,见(3-16)). 所以，(1-9)的平均$\leqslant C\cdot$(最大底)，是可见量. 于是全段误差有估计：

$$|\text{全段高}-\text{微分的和}|\leqslant C\cdot(\text{最大底}), \qquad (1\text{-}14)$$

其中常数见(3-17)，右边可算，无需 $\varepsilon-\delta$ 语言.

总之，认出了(1-9)的上界之后，便知(1-7)中全段误差当分割加密，会变得任意小. 如愿以偿!

上面的基本定理(1-11)先照着图形认出它的真面目，再用两行算术严格证明之(不至于似是而非)，无需实数(在情形 1)、无需$\varepsilon-\delta$语言(在情形(1-14))，连数学符号都不用，这可能是最小的代价，又是丰满的(恒等式，必要且充分条件). 多轻松!

到此，微积分的第二任务——图 1-6：曲线全段求高(微分的积分或基本定理)，计算相对误差的平均，也已经完成!

总之，参照图 1-2 和图 1-6，计算误差(1-4)和(1-8)，便轻松地到达微积分的彼岸：

总结　微分学就是曲线小段求高(图 1-2)及相对误差(1-4)，积分学就是曲线全段求高(图 1-6)及相对误差的平均，如

(1-14).

可见,相对误差(及其平均)是微积分的最重要概念,需要创造性的思考和严格的拷问.

如果微积分还要加上一个泰勒定理,那么后者的真面目又通过前者基本定理认出:

$$\text{泰勒定理}=\text{基本定理连用}, \tag{1-15}$$

即不停地使用基本定理来套,无需另找证明,多轻松!详见1.11节.

回顾正统教材,基本定理和泰勒定理的证明又长又难(前后几百页),记不住.记不住也好,这里只需要知道求高,自己来做,于是,只用两行算术就能证明基本定理(长了记不住).随之泰勒定理的真面目也一眼认出(不过是基本定理的连用)无需另找证明,只要讲几学时,多轻松!不过,这只是微积分的主部,或称初等微积分,而不是微积分的全部(后者包含微分方程,要用实数)!其实,一个人就怕容易的(初等微积分)不会做,那不应该;难的(高等微积分)不会做,倒无可非议.知足常乐!

注4 我们在中学学了那么多的定理,怎么没有见到基本定理呢?其实后者,或大学公式(1-11),由图1-6一眼认出,不过是中学正切公式一个挨一个垒起来.它给出了曲边三角的高和各点斜率之间的关系,是曲边正切公式.所幸的是,曲和直之间的差量:

$$\text{大学(曲)}-\text{中学(直)}=\text{误差}.$$

容易计算，当这误差≪1 能消失时，微元法才生效，才能认出

大学 = 中学的积分，

但大学和中学差别的最难还在于概念的转变：由不变(常量)到变量(特别地，由割线到切线)！

1.3　完整的微积分(初学略去)

可能有人会说，1.2 节给出的基本定理不完善，其中积分不是按一般的“黎曼和”来定义．那么能不能完善呢？回答是肯定的．

首先，图 1-1 切线不一定要取在起点，能取在任一中间点，见图 1-7．

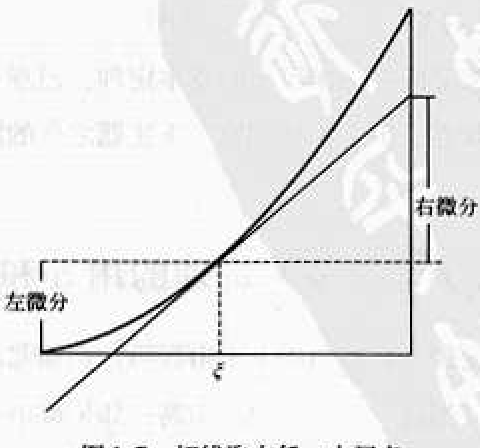


图 1-7　切线取在任一中间点

左右相对误差≪1．

这时微分设在中间点：(1-6)改变为

$$\text{微分}=\text{在}\,\xi\,\text{处斜率}\times\text{底},\tag{1-16}$$

其余不变，仍有(1-4)：ξ 左右相对误差≪1.

证明也不难，只要把中间点作为新分点，计算它的左右微分式，相加就是这一分点处的微分式，它的误差就是左右相对误差的平均，保持性质(1-4)．详见2.3节．所以(1-16)只是(1-6)的变形，没有实质差别.

那么，图1-6中切线不一定取在起点，也能取在各段的任一中间点(相应于黎曼和)，计算后仍有关系式(1-8)．但是这个中间点也可以看作新的起点，又变回(1-6)的情形，所以两者没有实质差别，后者只是前者的变形而已.

结论　当(1-6)改为(1-16)，仍有转换恒等式(1-8)，必要且充分条件(1-9)及基本定理(1-11)．在条件1)~4)下，也有估计(1-13)、(1-14)等，见3.2节、3.3节.

这才是完整无缺、无懈可击的基本定理，跟现有教材一致！

常人到此知足，但也不妨浏览一下主题之外的副产品.

1.4　副产品(基本定理的用处和无能)

回到基本定理．它真的用于求山高吗？正如北京大学刘嘉荃所说，没有人会用这个基本定理求山高：如果说山坡上的各点斜率都能计算(只涉及一点附近的性质)，相应的底长又怎么计算呢？这几乎不可能．但是基本定理首先是指出山高和斜率、积分和微分之间的关系，当然也能用来计算曲边图形的面积，以及套

出泰勒逼近及微分学各定理. 此外, 微元法能计算弧长, 甚至解读托尔斯泰的小说. 这些只是作为基本定理的副产品, 无需另找方法. 可见, 反复使用一个基本定理(见 1.2 节)或微元法的思想, 便能做出很多事. 但是, 必须指出, 基本定理也非万能, 第 5 章(或 1.12 节)将看到, 它就解不了应用数学中最重要的微分方程(5-3).

下面的内容在初读时不妨略去, 有耐心的读者可以继续往下读.

1.5　从求高转到求面积
(基本定理的另一解读)

现在由求高转到求面积. 正统教材把基本定理解读为求面积(间接法, 凡人想不到), 本书解读为求高(直接法, 凡人也想得到). 两者相通吗? 幸亏能够转换: 测量山坡图(上面的曲线)所记录下的斜率图(下面的曲线,或称山影)围成的面积, 能够认出它即是山坡图自身的求高, 看图自明(图 1-8).

图 1-8 曾在 1999 年出现在我的《画中漫游微积分》一书封面上, 2003 年国外一书用了同样的图, 并说 "这大概是单一最重要的图, 顶得上一千个符号和方程, 将积分的实质压缩在一张快照之中".

所以积分还有另一解读: 曲线下求面积, 返回到原山坡求高, 多轻松!

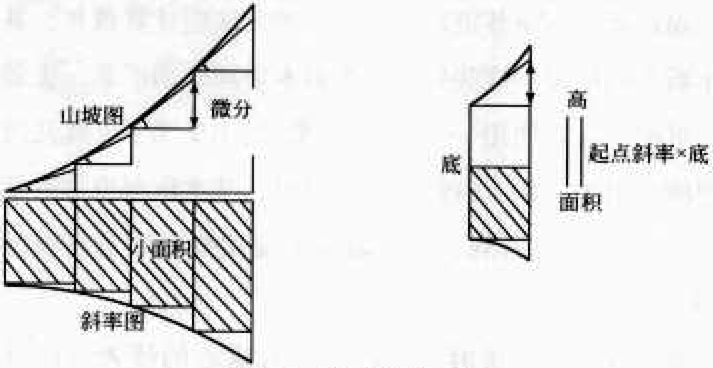


图 1-8 斜率图求面积

斜率图的总面积 = 山坡图的总高.

从物理看，高和面积的量纲不同，但是从数学看，数值相同.

1.6 弧长与斜率(照猫画虎)

为了加深对微元法的掌握，再举一例：测量一条“可求长曲线”的弧长要用斜率.

现在再次采用微元法：照猫画虎，复述图 1-6 的做法，把全弧分割成许多小弧，再把小弧换成起点处的切线(图 1-9).

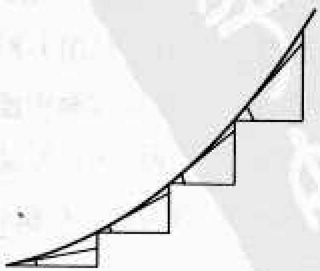


图 1-9 弧长测量

小弧长≈切线长.

现在，把小弧长换成在起点处的切线长，后者称为弧微分，

也要用斜率来算．由勾股公式：

$$弧微分 = \sqrt{(底)^2 + (微分)^2} = \sqrt{1 + (斜率)^2}\ (底). \tag{1-17}$$

凭感觉，

$$全弧长 = 弧微分的积分. \tag{1-18}$$

由图 1-9 一眼认出，它不过是中学勾股公式一个挨一个垒起来，是曲边勾股公式．严格说，曲和直之间的差别要经过计算(第 5 章附录将证明这个积分的存在性)，微元法才生效，两者才统一．正如勾股公式首先是指出三边之间的关系，当然也能用来计算斜边的长．同理，这个弧长公式首先也是指出不同量之间的关系，当然也能用来计算圆的周长等．阿诺德(2006)给出一例，指出一个标准正弦曲线的长度只比横轴的长度 2π 大约增加 20%（图1-10）．

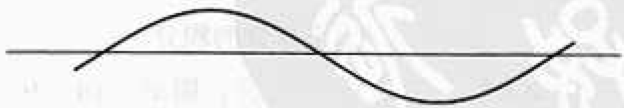


图 1-10　正弦曲线之长

例如，5 米的横轴大约相当于 6 米的标准正弦曲线．直观只能说曲线比横轴长，只有弧长公式才能定量．这里真正用到了斜率．

1.7　微积分一张图

到此，微积分是一个测量学，测斜率、弧长、高和面积．它们能在图 1-8 或图 1-9 中认出．

1.8 人口预测

微积分有什么用呢？有许多几何、力学问题要用微积分，这些是无可置疑的，不过百姓对此未必会被触动．这里只举一个有切身感受的例子，即中国人口普查．中国人口普查如用微积分，能把几亿户一年的工作量化约为一人几分钟．多轻松！见5.1.2小节．这就是微积分的威力！但是人口问题不能由1.2节的基本定理来解，需要引入新的解，需要实数．

1.9 《战争与和平》的解读

现在转到托尔斯泰(1987)的历史学基本定理：

总历史＝历史微分的积分．

不能用一个历史人物的活动来表达(注：附录“四”中讲的中值定理又说，理论上存在某个救世主，就是找不到)．对于数学工作者，这一解读还不能满意．但是，基本定理提供了思维的微元法．这样的例子很多，包括管理学的原理也是先做好微分(细节)后再积分．详见本章附录之十．

以下各节初读略去．

1.10 民间和主流

以上只用白话说事，没有数学符号和形式推理，称为民间微

积分，可能被讥讽为“胡来”．其实，白话说事也能完成基本定理等的精确研究，已见于1.1～1.3节．历史上阿基米德的数学著作，据说没有一条公式，都是用文字和图形．只是文字和图形难以推导更一般、更复杂的泰勒定理，只好采用函数体（称为主流微积分），见第4章．先睹为快，我们摘出其中一段．

首先，基本定理的白话体能够翻译成函数体．用笛卡儿坐标：设山底的点 x 由 a 跑到 b，山坡表述为高 $f(x)$，山坡上的斜率（称一阶导函数）表述为 $f'(x)$，微分表述为 $f'(x)h$（图1-11）．

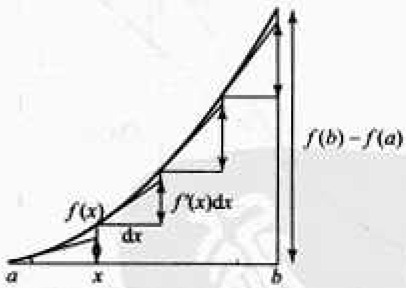


图1-12　微分的函数表示

图1-12中起点 x 的“附近” h 是变量，测量误差表述为

$$f(x+h)-f(x)-f'(x)h=\epsilon(x,h)h, \tag{1-19}$$

其中相对误差 $\epsilon(x,h)\ll 1$（即绝对值可以任意小，当 h 变小）．h 又记为 dx（点 x 的附近，或增量）．这是不是难以接受呢？法国数学课程标准(2001)高三部分就是通过(1-19)来刻画微分的！

山坡的全段高表述为 $f(b)-f(a)$，那么基本定理(1-11)表述为图1-12．

当 f 平均可微（即可积条件(1-9)）或一致可微（定义为(1-12)）

$$f(b)-f(a)=\int_a^b f'(x)\,dx$$

(其中变量 x 能任意取，如取 s). 这就是在笛卡儿坐标里由导函数 $f'(x)$ 给出的曲线 f 的高度(见图1-12. 伏笔：这就是微分方程 $f'=g$，由 g 给出 f 自身). 所以，1.1 节、1.2 节的白话微积分跟正统教材相比照，完全一致，没有不准确或不严格之处，可以放心！

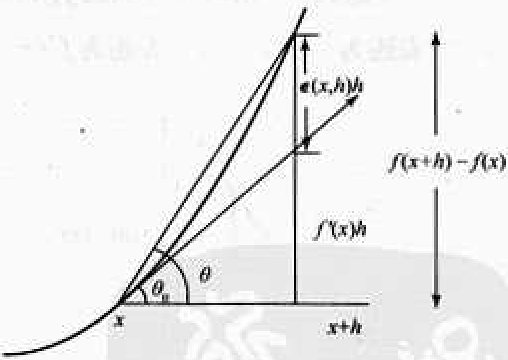


图 1-11 基本定理一眼认出(图形和函数相结合)

特别地，对于 $f(t)=t^{n+1}$ ($n=0, \pm 1, \pm 2, \cdots$)，有导函数 $f'(t)=(n+1)t^n$ 及积分

$$\int_0^x t^n \mathrm{d}t = \frac{x^{n+1}}{n+1}, \quad n \neq -1. \tag{1-20}$$

但是，当 $n=-1$，即被积函数为 $1/t$ 时，不能应用基本定理求积分，能补救吗？见下文 1.12 节.

1.11 泰勒定理

现在由基本定理转到泰勒定理. 1.1 节把曲线小段用切线

(一次函数)代替，本节则用二次或更高次的函数代替. 然后，小段高用高次曲线的高代替，误差越来越小. 幸亏，计算并非越来越难，只不过是基本定理的连用几次. 事实上，凭借基本定理做平台，一层层叠起来，无需另行推导(分部积分)就能垒出一个泰勒定理. 首先，基本定理本身就给出一个常数逼近：令 h 为变量，在分点 x 处，

$$f(x+h)-f(x)=\int_x^{x+h} f'(s)\,\mathrm{d}s :\approx h \underset{[x,x+h]}{\text{upper}}|f'|$$

(副产品包括微分学诸定理：$f'\equiv 0\Rightarrow f=C$, $f'>0\Rightarrow f\uparrow$)，这里采用了缩写 $a:\approx b$ 表示 $|a|\leqslant b$，upper 表示上界. 其次，微分(1-19)就是一个一次逼近：在分点 x 处，

$$f(x+h)-f(x)-hf'(x)\quad (\text{其中} f'(x) \text{是常数})$$

$$=\int_x^{x+h}[f'(s_2)-f'(x)]\,\mathrm{d}s_2=\int_x^{x+h}\int_x^{s_2} f''(s_1)\,\mathrm{d}s_1\mathrm{d}s_2$$

$$:\approx \underset{[x,x+h]}{\text{upper}}|f''|\int_x^{x+h}\int_x^{s_2}\mathrm{d}s_1\mathrm{d}s_2=\frac{h^2}{2}\underset{[x,x+h]}{\text{upper}}|f''|.$$

这里的诀窍(如果有)就是每逢相减变成一个积分.

到此已经认出泰勒定理的真面目，它不过是基本定理连用两次. 多轻松!

注 5. 现有教材用的分部积分，难记. 这里我们只记住基本定理. 急中生智，就直接用基本定理，连用几次就套出来了，多轻松!

注 6　副产品：相对误差 $:\approx Ch$，与点 x 无关，并随着 h 一起变小.

下面知道该怎么做了! 继续延伸，有二次逼近：在起点 x 处，

$$f(x+h)-f(x)-hf'(x)-\frac{h^2}{2}f''(x) \quad (\text{其中 } f''(x) \text{ 是常数})$$

$$=\int_x^{x+h}\int_x^{s_3}[f''(s_2)-f''(x)]\,\mathrm{d}s_2\mathrm{d}s_3$$

$$=\int_x^{x+h}\int_x^{s_3}\int_x^{s_2}f'''(s_1)\,\mathrm{d}s_1\mathrm{d}s_2\mathrm{d}s_3$$

$$:\approx \operatorname*{upper}_{[x,x+h]}|f'''|\int_x^{x+h}\int_x^{s_3}\int_x^{s_2}\mathrm{d}s_1\mathrm{d}s_2\mathrm{d}s_3$$

$$=\frac{h^3}{3!}\operatorname*{upper}_{[x,x+h]}|f'''|.$$

三次逼近：在起点 x 处，

$$f(x+h)-f(x)-hf'(x)-\frac{h^2}{2}f''(x)-\frac{h^3}{3!}f'''(x)$$

$$:\approx \frac{h^4}{4!}\operatorname*{upper}_{[x,x+h]}|f^{(4)}|.$$

高次逼近：在起点 x 处，

$$f(x+h)-f(x)-hf'(x)-\frac{h^2}{2}f''(x)-\cdots-\frac{h^n}{n!}f^{(n)}(x)$$

$$:\approx \frac{h^{n+1}}{(n+1)!}\operatorname*{upper}_{[x,x+h]}|f^{(n+1)}|. \tag{1-21}$$

这里用到高阶导函数（$f'', f''', \cdots, f^{(n+1)}$）和累次积分，以及一个控制式：$|g|\leqslant C$，则 $\left|\int g(s)\,\mathrm{d}s\right|\leqslant C\left(\int \mathrm{d}s\right)$．（1-21）说明了一个充分可微的函数很像多项式：

泰勒定理

可微函数 ≈ 多项式.

其中的差量为高阶量.

到此已经一眼认出泰勒定理的真面目，它正是(1-15)所说的一连串基本定理垒起来，即基本定理的连用，多轻松！正统教材靠分部积分的技巧，难记．学生更欢迎机械的操作．这就是基本功；假如使用基本定理，便能一眼认出泰勒定理，何必舍近求远，付出分部积分的代价呢？假如用一行(或一页)能套出，何必用两行(或两页)？熟知为每个学生省一行(或一页)，年年省下来，会有多少学生受益！

到此，微积分又一道难关——泰勒定理，无需动用实数，也闯过去了．真是有惊无险！下一道难关，最实际的微分方程(5-3)的求解，躲不开实数，见 1.12 节．就像碰到方程 $x^2=2$，没有实数就没有解．

1.12　微分方程的求解(有赖于实数)

按 M. Livshits(2008)的路线，先对数后指数．积分表(1-20)漏了 $n=-1$：先记为自然对数函数

$$\ln(x):=\int_1^x \frac{1}{t}\mathrm{d}t. \tag{1-22}$$

我们得不到一个显式的原函数 $f(t)$，使得 $f'(t)=1/t$．不能利用基本定理(1-11)，于是只好直接计算这个积分的左(分点)右(分点)黎曼和(图 1-13、图 1-14)．

可以验证（详见 3.5 节）：

右和 < 左和，　左和 − 右和 ≈ 分割长，

以及左右和的单调性：加密后，右和变大，左和变小．它们彼此

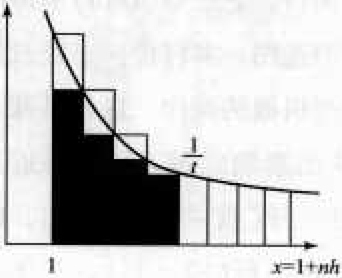


图 1-13　左和与右和

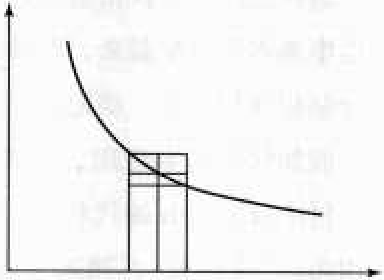


图 1-14　左右和的单调性

靠近，且不会越过．所以，其中夹着一个数(实数)．还可以证明这个数不依赖于加密方式，它是惟一的，把它定义为积分或自然对数(1-22)．结果，这个实数为积分表(1-20)填补了漏洞．可谓锦上添花!

容易验证，对数的导数

$$(\ln x)'=\frac{1}{x}.$$

把对数函数反过来，就可以定义指数函数：

$$\ln(e^x)=x,$$

它的导数为

$$(\ln(e^x))'=\frac{(e^x)'}{e^x}=1,$$

所以 e^x 的导数为

$$(e^x)'=e^x.$$

这就得到微分方程 $f'(x)=f(x)$ 的解．所以，“副产品” e^x 比原发品 $\ln x$ 还要重要得多!

如果对数是锦上添花，指数便是雪中送炭!

1.13　筋　骨

本书的快餐又分为早餐(第1章)、午餐(第2~5章)和晚餐(第6、7章及另册).

微积分的筋骨——两大概念(微分和积分)和两大定理(基本定理和泰勒定理),是以后应用的平台,我们用几页(两张图、两行算术,不用实数)把它搭建起来.这是早餐或初等微积分(1.1~1.3节,1.11节).这样,应用科学就能和数学同步教学.有了这个筋骨或平台,对一般人或文科的学生也就够了(解读托尔斯泰的小说,以及简单的物理应用).要割爱,一个人一生能留住这些也就知足了(少则留,多则忘);容易的要会,难的忘了无所谓!接下来,对理科,要更上一层楼,需要计算微分方程(这时躲不开实数了),这是午餐或中等微积分(第2~5章).更专业一些,到第二学期(第6、7章及另册),再吃晚餐:完善理论(高等微积分,包括实数、ε-δ语言、连续函数、中值定理、一般的原函数的存在性).这有点奢侈,多数人受不了,但为满足有余力或考研的学生.

本书将微积分劈成两半,以实数划线,前半(两个定理)不用实数;后半(一个方程)才用实数,这时必须学习陶哲轩(2008)的实数论.这与正统的教材是两种写作方式.正统教材

是由一般再到特殊①，即先学一般理论(高等微积分)，急需的定理放在最后，可惜物理、应用等不及了；而本书不顾一般，破门而人，直取基本定理(便有泰勒定理)，不计其余，代价最小．但这种写作方式自从在光明日报和人民日报(林群,1997)刊登后，十年来只有极少数大学或学院采用(如深圳大学(1998)、广西师范大学(1999)、河北大学(2002)、大连理工大学(2006)、滨州学院(2006)、中央财经大学(2007))，更多的学校只是用了求高图(图 1-6 或图 1-8)(例如复旦大学(1999 年以来)，Wiley Publishing (2003)和清华大学(2006))．也就是说，绝大部分专家并不认可我的这种写法！尽管如此，还是有一家出版社出版了我按此方式写成的书（可惜用的是英文，表达不当)．现在这本书用的中文，如鱼得水，并经过了装饰和补充．我期望在教学实践中会逐渐得到更多人的理解．

令人鼓舞的是，国内有张景中(2008)完整地建立起基于初等数学的微积分系统，避开了导数的极限概念和黎曼积分的定义，将微积分初等化！

国外至少有拉克斯等(1980)，G. Strang(1991)，K. Dovermann(1999)和 H. Karcher(2002)对微积分进行改造．更接近我们的，有 M. Livshits (2004)，他也只讲微分和积分，但有自己的特色，自圆其说．特别是，他对自然指数引入的路线，改变了本书对微分方程一章的写法．他对基本定理充分条件的证明避开了中值定理，

① 据 M. Livshits (2008) 介绍，Vladimir Arnold 有一“最小一般性原理”．据此，每一想法先由最简单的情形入手，之后这种方法再被推广到复杂情形．

也为本书（3.3 节）所采用. 此外，陈掌星、鄂维南、沈善普、吴从忻、A. Leger 和 H. Brunner 也有同样的想法、目标和兴趣. 特别是 M. Campbell 准备改编本书，这说明微积分的简化到时候了.

1.14　血和肉(应用和练习)

本书只提供微积分的筋骨. 但是光有筋骨还不够，微积分的生命——血和肉，在于应用和练习. 学数学不做练习就像游泳不下水. 本书只挑出周围世界发生的例子：除了几何例子，有中国人口预测以及解读托尔斯泰小说，帮助一般人或文科的学生了解微积分的意义和威力. 但是对于理科的学生还远远不够，必须配合其他教材，例如，华罗庚(1963 年以来)，G. Strang(1991)，Y. Zeldovich-I. Yaglom(1987). 这些教材备有各种应用和练习，各位教师根据自己的专业需要和学生水平从中选择，加入所需，编写成符合本专业和学生水平的教材，这里就不包办，只是留给教师一个空间，往我们搭建的筋骨里添加血和肉，或者说，我们只是把大楼盖好，内部装修则靠教师自己.

总之，本书只是母本(不是全本)，只有微积分的筋骨，不含血和肉，只安排极少量应用和练习，需要教师自己从 Y. Zeldovich-I. Yaglom 的书中(网络上)去找，再编写成子本.

附录　登山对话

这个附录对初学者起着引路作用，对已学者起着复习作用.

这里微积分被叙述成登山时对斜率(陡度)和高度的测量. 下面是学生和老师的对话.

一、山坡走势和斜率

他们往上爬，达到一个平台后休息(图 1-15). 然后继续往上爬.

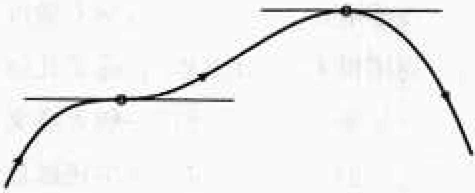


图 1-15 光滑山坡

生：那么这个平台是不是微分学中的稳定点?

师：你认出来了. 到达平台之前山坡在上升，过了平台继续上升，这就叫做稳定点.

他们爬到第二个平台，然后下山.

生：那么这个平台是不是微分学中的局部极大?

师：你认出来了. 到达平台之前山坡在上升，过了平台开始下降，这是一个高峰，又称局部极大.

生：到此我们有了山坡的上升、下降、稳定点、极值(极大或极小)点，这些感觉常人都有，不需要数学呀.

师：小学有了自然数，把常人都有的“多或少”的感觉量化为数字，以识别多或少的程度. 微分学发明了斜率这个数，把山坡是怎么走的化为斜率的符号(正、负、零、变号)和大小:

当沿山坡往前走时上升或下降，升降得快或慢，在什么地方达到稳定或取到极值，以及山坡是怎么弯曲的.

生：斜率对于登山竟然如此基本，能不能将斜率这个概念说明白？

二、斜率和切线

师：如果山坡是直的，斜率就是测量这山坡的方向(图1-16).

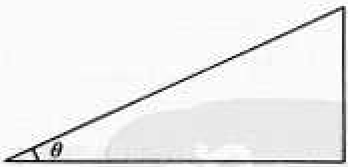


图 1-16　直的山坡

如果山坡是一条光滑曲线，一点在上面运动着，于是它在每一点有一个方向，同时连续地改变它的方向(图 1-17).

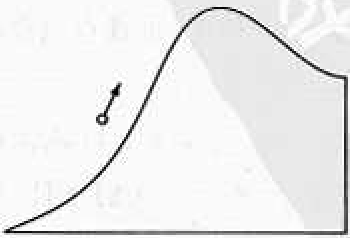


图 1-17　曲的山坡

生：感觉会有这样的方向，但是这个方向用什么来决定呢？

师：由切线来定，这是微分学的诀窍. 切线是什么呢？现有

教材上都有，为了在点 t 处决定切线，在曲线上取与点 t 不同的点 s，做割线 ts. 然后让 s 沿着曲线去接近点 t. 如果这时割线 ts 趋向一个极限位置(这是一个假设)，则这个极限位置的直线就叫做曲线在点 t 处的切线，见图 1-18.

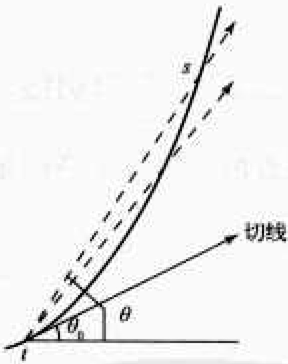


图 1-18 曲线在点 t 处的切线

显然，切线角(θ_0)应该等于割线动角(θ)的极限.

生：感觉会有这样的极限. 但是割线的极限位置或切线一定存在吗?

师：不一定. 如果(或假设)曲线点点存在切线，简称这曲线点点"可微"或处处"光滑".

生：为什么要用切线? 用其他直线(割线)又会怎样呢?

师：切线有特殊的性质. 让我们来计算由切线到曲线的"距离"，或起点附近(图 1-19)，

测量误差 = 小段高 - 切线高，

切线高 = 斜率 × 底.

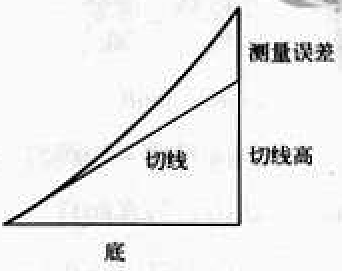


图 1-19　求小段高

生：图 1-19 上看得见，无须花脑筋计算，这个测量误差随着曲线小段变小会变小.

师：但是对于切线来说，更有用的是"相对误差"（关键词）不是误差自身，或测量误差与所取曲线小段(或底)之比：

$$相对误差:=\frac{测量误差}{底}.$$

生：它又是什么呢？看得见吗？

师：由图 1-19 切线的构造，看到起点附近(图 1-20)，

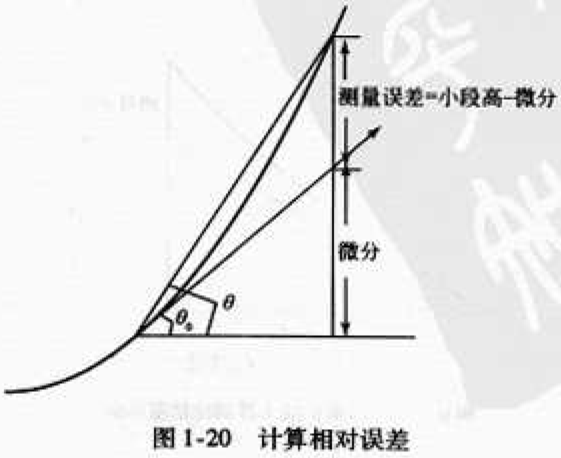


图 1-20　计算相对误差

$$\frac{\text{测量误差}}{\text{底}}=\frac{\text{小段高}}{\text{底}}-\frac{\text{微分}}{\text{底}}$$

$$=\tan\theta-\tan\theta_0$$

$$=\text{割线斜率}-\text{切线斜率}.$$

更清楚些，通过三角关系认出它的真面目

$$\text{相对误差}\leqslant C|\theta-\theta_0|$$

右端已是可见的变量(其中 θ 是割线动角，C 是常数)．看得见它随着割线趋向切线(或底变小)会变得任意小.

生：以上是图形或感觉，怎样定义切线呢?

师：先感觉后严格．以上是对切线或斜率的感觉，不能算数．反之，今后就把起点附近“相对误差≪1”(即绝对值小于任一指定的数)作为切线的定义性质(惟一性见第 1 章注 2)．这也是本书惟一要用的定义！缺它不行，初学者又最难接受“相对误差”这个概念，总觉得别扭，但是法国数学课程标准(2001)高三部分都这么讲，必须有信心！忘了怎么办？看图自明(图 1-21).

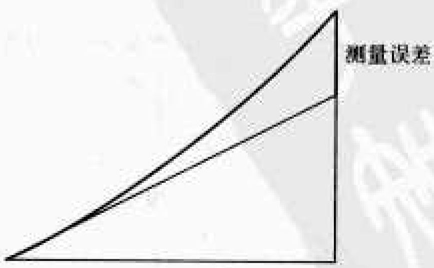


图 1-21　切线性质

测量误差与所取曲线小段相比是很小的.

最好去看 Arnold(1995)的动画：如果底很小，测量误差便非常小，这就是为什么相对误差会很小.

生：那么，其他直线(割线)的相对误差也能变小吗?

师：不能．切线的相对误差趋于零，但割线的相对误差趋于非零常数．看图自明(图 1-22).

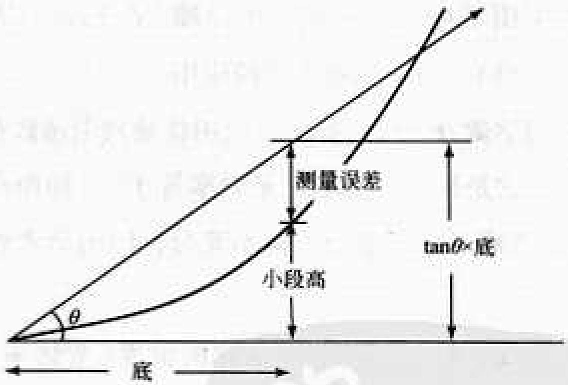


图 1-22　割线的相对误差

对任意割线如果斜率为 $\tan\theta$，则有

$$\text{相对误差：}=\frac{\text{测量误差}}{\text{底}}=\frac{\tan\theta\times\text{底}}{\text{底}}-\frac{\text{小段高}}{\text{底}}$$

$$\to\tan\theta-\tan\theta_0=\text{常数}，\quad \text{当底}\to 0\text{ 时}.$$

生：明白了，只有切线的测量误差才是最小(最好的逼近)，这就是为什么微分学把曲线小段换成了切线(最好的代替)，而不是割线.

师：中学只是对于直线才有方向的意义，现在被推广到曲线上：每一个点的曲线方向就由这一点处的最贴近的那条直线——切线(最好的逼近)来决定，从而获得了原来没有的意义．这是

非常明智的！

生：这应该是大学学到的第一个新东西！

三、斜率的普遍应用

师：斜率用于登山，只是一个比喻（首先见于法国哲学家C. Bruter(1973)）. 斜率有更普遍的应用.

生：我们学微分学时，有一个应用就是判别函数在什么地方能达到极值，它是说极值点必定是斜率等于零（即曲线的切线是水平的点），变成了方程求根. 但不充分，因为后者也可能是稳定点.

师：一般说，斜率用来度量变化的世界（变化率），包括经济的走势，升或降，快或慢，或稳定. 例如，俄罗斯数学家阿诺德(2006)说到用斜率的符号度量经济的走势.

生：第一段凭着斜率这个数，已经能够量化地叙述山坡（或变化世界）的走势，山坡一点附近的性质. 斜率还有别的能耐吗？

四、山高和斜率：非构造的正切公式（中值定理）

师：斜率也能用来测量山高. 中学测量树高，便是通过一个斜率，见图1-23，不必砍树. 只是山坡是弯的，见图1-24.

山坡上不同的点，会有不同的斜率（为好画，下面只取山坡

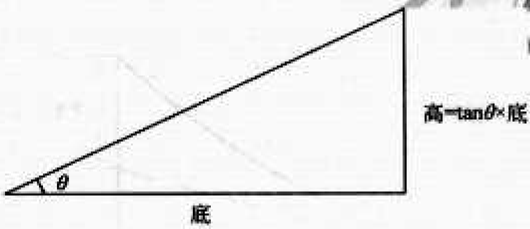


图 1-23　测树高

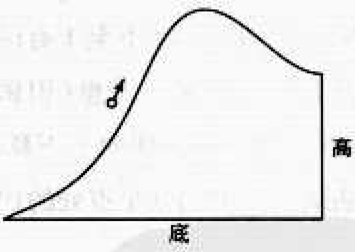


图 1-24　测山高

的一段)，见图 1-25.

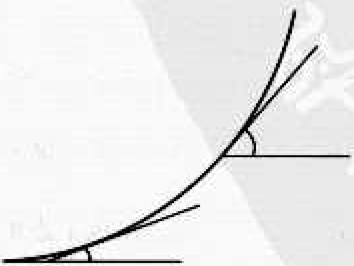


图 1-25　斜率在变

生：那么要取哪一个点的斜率来测量山高呢?

师：这是微分学的基本问题. 如果随便取(如在起点)，可

能有大误差(图 1-26).

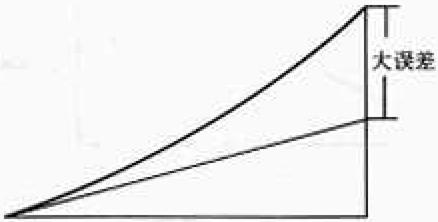


图 1-26　任取斜率的后果

生：那么怎么取这个斜率，才不至于有误差呢?

师：感觉会有这样一个斜率. 设想(但做不到)把山脚和山顶连成一条线(虚线)，然后平行地向上下移动，则在连线脱离山坡的时刻，它就占有这山坡的某个点处的切线位置(图 1-27).

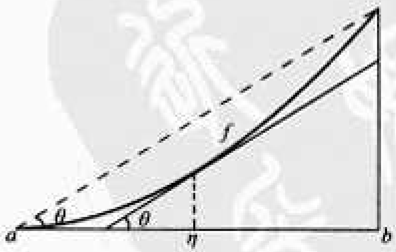


图 1-27　中值定理的几何说明

这就是说，在山坡上存在一个点，在这个点的切线角与连线角相同. 于是，这个切线角的正切(k = 斜率($\tan\theta$))乘上山的底长便是山高，$f(b)-f(a)=k(b-a)$. 这就是曲边三角形的正切定理，又称中值定理.

生：认出来了，原来大学中值定理不过是中学正切定理的推

广，两者是统一的．把中学(直)搬到大学(曲)，没有一点误差，真聪明！

师：但是这个点取在哪儿呢？我们是近视眼(或蚂蚁)，只看到一点附近的性质(斜率)，看不到远方，无法去连接山脚山顶这条连线，也不会将它平行地移到山坡的某个点上．所以这条切线取在山坡上的哪个点是未知的！这是一大缺点，这就像买彩票一样，明知头彩是存在的，但是能摸出哪个号预先未知．而且，中值定理的证明要用复杂的理论．

生：太可惜了，中值或 k 存在但又找不到．能不用中值定理吗？

师：M. Livshits(2008)避开中值定理，照样证得很多有用的定理(如 3.3 节)，但要用一点实数．而张景中(2008)基于一个非常显然的思想：平均速度一定介在“瞬时速度”之间．所以一定存在这样的点 η^1，η^2，使得

$$f'(\eta^1)\leqslant \frac{f(b)-f(a)}{b-a}\leqslant f'(\eta^2),$$

这不是等式，可以代替中值定理(等式)．

五、构造性的正切公式：基本定理

师：中值定理最聪明，但是不能用作精确计算(因为中值是未知的)．牛顿和莱布尼茨则用最笨的办法，即所谓的微元法：化整为零，分而治之，由零到整．

生：怎么化整为零呢？

师：第一步化整为零，就是将山坡全段缩短成小段(图1-28).

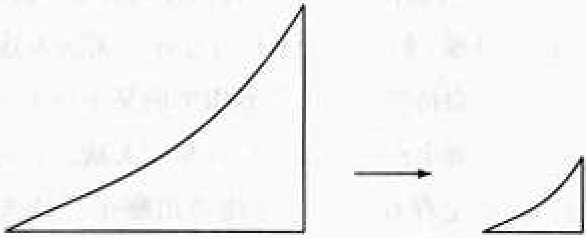


图1-28　曲边三角形被缩小

生：但是曲的还是曲的，这一步还看不出好处！

师：但是曲线小段的每一点斜率差不多相同．第二步分而治之就是选取其中任一点，如起点，以此点处的切线斜率作为曲线小段的斜率(图1-29).



图1-29　小段高的计算

据此计算出切线在起点附近的高，称为微分高或简称

$$微分 = 斜率 \times 底.$$

生：但是它并非曲线小段的真高，含有(图1-30)：

$$测量误差 = 真高 - 微分.$$

师：上面“二”中讲切线时已经计算过这个误差，它非常小，以至于和所取曲线小段(或底)相比，更有用地，

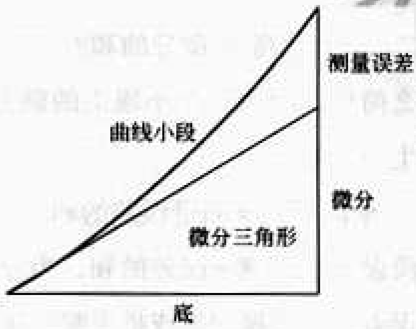


图1-30　曲边三角形和微分三角形

$$相对误差:=\frac{测量误差}{底}$$

也很小. 这就达到化整为零，分而治之，完成了曲线小段求高(微分，最好的逼近).

生：那么，第三步如何由零到整，延拓到全段上?

师：它就是将分段小段上的微分(中学直边三角形的高)加起来(图1-31).

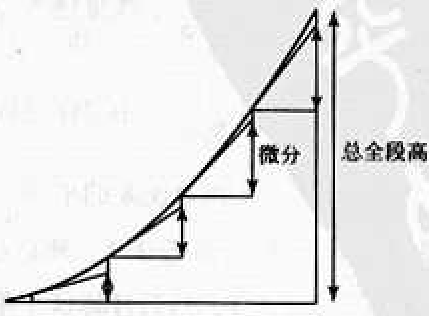


图1-31　总全段高和微分之和

生：那么应该有

$$全段高 \approx 微分的和?$$

师：没有那么简单，因为各个小段上的微分都有测量误差，加起来之后会产生一个

$$全段误差 = 测量误差的和.$$

生：这个全段误差，或测量误差的和，当分割加密，项数会越来越多，那么误差会不会被积累或扩大呢？还是反而会变小？

师：这是最令人担心的问题，究竟怎样，不能凭感觉，有待进一步计算．只有这个和不因加密（和项增加）被放大，反能小于任一指定的数，第三步才有效．

生：那么这个和怎样计算，以便判断会不会被放大呢？

师：跟上面"二"中一样，还是来计算"相对误差"，但是现在指全段误差和全段底（或总底）之比

$$\begin{aligned}\frac{全段误差}{总底} &= \frac{(测量误差)的和}{底的和}\\ &= \left(\frac{底}{总底}\cdot\frac{测量误差}{底}\right)的和\\ &= \left(\frac{底}{总底}\cdot 相对误差\right)的和\\ &:= 相对误差的平均\end{aligned}$$

（其中各项系数为正，而且加起来等于1，所以是加权平均．但当各底相等时，便是算术平均），然后计算这个平均．

生：这个平均又是什么呢？看得见吗？既然已知各段相对误差都能任意小（因为曲线各点可微），而它们的平均又不会超过

其中某一个相对误差，那么这个平均就应该任意小吧?

师：没有那么简单．我们担心的是，当网格加密，各段相对误差会不同程度地变小，不见得一致．所以，什么时候平均小，需要交代清楚．

生：我已经注意到图 1-31，那里的相对误差(或测量误差)在各段的分布不见得均匀．

师：如果相对误差在各段的分布均匀(称曲线“一致可微”)，即相对误差能一致变小，不依赖于分点的位置，那么能有平均变小．幸亏这个一致变小的假设容易满足(相当于曲线“连续可微”)，这时，当网格加密，全段误差才能小于一个指定的数，便有可积性以及基本定理：全段高 = 微分的积分．这里把“离散”的和写成“连续”的积分．这样才达到严格化，不至于导致误解，或模棱两可、似是而非．

生：没想到基本定理的证明这么简短(不到一页)，只用到平均变小的条件．正统教材必须用中值定理(其证明巧妙且长，参见上面“四”)．

师：不仅简短，实际上找到了基本定理或可积的充要条件：平均变小．数学追求完美(充要条件)．上面说的“连续可微”条件只是特殊情形，充分非必要．

生：许多人敬畏数学，是因为数学需要高技巧．老师都要求多练技巧，做奥赛题．可是你所说的微积分基本定理，没有什么技巧，只用到平均的思想，后者是物理实验常用的办法，不是特殊的概念．

师：不能盲目追求技巧. 一般说更大的技巧(或工夫)在于能把通过繁杂计算才能达成的理论认识变得一眼认出、一举得证. 数学家华罗庚就是这方面的高手. 当过北京大学校长的丁石孙说："华罗庚当时非常得意，他新完成一个工作，就是证明了所谓射影几何的基本定理. 他说他证明这条定理，只用了初等数学的方法，只一页纸就写出来了. 过去人们用了很多高等数学的知识，证出的只是特例. 后来我才知道，华罗庚只用了很简单的一个公式就证出来了. 他的证明，现在一般称作华氏定理."《科学时报》也报道过数学家张寿武谈数学心得："我特别喜欢做数学的过程，坐在那里慢慢地思考、重新规划，把一个非常复杂的问题弄成一个很小的问题. 我觉得数学最妙的地方是：正确是基于简单的理由，而不是复杂的理由. 数学与科学和文学一样，能够留下来的东西都是最简单的."

生：这太费时间，太慢，原地踏步怎么赶上前沿？要学得快，只能狼吞虎咽，先灌进去，赶快跑到前沿去.

师：这是一对矛盾. 不过要选好书，避开长篇阔论，很快入门.

六、大学 − 中学 = ?

师：基本定理的真面目由图 1-6 一眼认出，不过是中学(直边)正切公式一个挨一个垒起来. 它给出了曲边三角的高和各点斜率之间的关系，是曲边正切公式. 所以大学的真面目由中学认

出．而它们之间的差量也能计算，当它们的差量≪1，能消失、微元法才有效，才能认出

$$大学公式 = 中学公式的积分.$$

反之，也能认出中学 = 大学的微分．

生：原先以为大学和中学完全不同，是全新的东西．没有想到能把大学和中学统一起来，温故而知新．原来大学的知识，认出后，不过是中学所学的推广啊．要学会轻松地学习．很多同学由于没有认出这一点，不把大学跟中学衔接，难以适应而失去信心，很可惜．

师：虽然认识大学是中学的推广，它们之间还是有质的差别（曲与直、变与不变、切线与割线），需要计算这个差量并消除差量，两者才统一．

生：大学跟中学对比，有同有异．

师：这里有一个学习方法的问题，数学家华罗庚提倡学习要做减法，去同存异，由厚到薄．所以，要分离出哪些是中学有过的？大学(曲或变)减去中学(直或不变)剩多少？所剩部分：

$$大学（曲）-中学（直）=误差$$

的计算，才是大学的工夫．

生：下意识地做这个分离工作和减法，认出相同和不同，是对大学生 4 年的基本要求！我这个大四的学生，前几年就只做加法，越念越厚，念完大学忘了中学，念完后面忘了前面，就因为没有做分离和减法，没有认出相同和不同．

师：当你们学到第 7 章时，会进一步加强大学跟中学的统

一，前者往往是后者的推广，甚至是照猫画虎.

七、其他课程

生：改写微积分，由厚到薄，保证简单和正确(伍鸿熙这么说)，让我们从痛苦的学习中解放出来. 能不能也改写线性代数和概率论呢？

师：我还不能. 改写需要灵感，捡到酵母才能发酵，就像写小说，没有灵感或酵母，怎么发酵呢？但我推荐线性代数的两位改革者，一位是美国的数学教授 G. Strang（1991），另一位是中国的数学教授李尚志. 对于概率论，我的学生沈小平，美国应用数学教授，推荐了 G. Papanicolaou(1973)的普及写作，将概率论跟二阶微分方程联系起来，见 5.3 节的后记.

八、专业研究

生：以上讲怎么学微积分，专业研究是不是也这样呢？

师：原则一样. 一位德国同行对专业演讲提出四步曲：第一步中学生能懂，第二步大学生能懂，第三步研究生和专家才能懂(那是专业)，第四步只有你自己懂(那是最新成果)，或连你自己也不懂(那是未来发展). 所以专业研究也要下分离的工夫，把专业中最原始的部分认出来，分出层次，认出哪些是中学、大学或专业课中有过的，哪些才是你的新成果以及下一步的发展.

九、普　及

生：还有一个问题，数学家华罗庚专业研究那么深，为什么最后转向推广优选法(0.618，黄金分割数)呢?

师：据数理统计学家张里千说：华罗庚说数学太难了，所以去抓 0.618. 黄金分割的这个数古代就有，不好说是创新，可是他讲得最好，讲成世界第一，就能感天动地！一次我乘出租车，当司机得知我也搞数学，蹦出一句："哦，0.618"！还听我的老师关肇直说，毛泽东见到华罗庚(大概 1958 年)，说读了他在光明日报写的水坝数学，觉得很好！华罗庚讲的东西，从百姓到领袖都听得进去. 这就是他的数学工夫或融会贯通. 还有数学家吴文俊也有这个本领.

生：现在数学家中，普及做的最有影响的是张景中吧？他曾是全国科普作家协会的会长.

师：张景中除了对微积分的初等化做出独到的贡献，还敢碰中学几何和三角的问题，卓有成果！他也是影响广泛的人物.

十、微元法的其他应用

生：没有人会用基本定理来测量山高，即使山坡上的斜率点点都能测量(只涉及一点附近的性质). 相应的底又怎么测量呢? 这几乎不可能. 所以基本定理会有什么用呢?

师：基本定理首先在于指出山高和斜率，积分和微分之间的关系，但是也能用来计算曲边图形的面积和弧长等，已见于第1章．此外，微元法有种种用处，甚至解读托尔斯泰的小说．你若有耐心，请继续看下去．

生：据文学界人士说，托尔斯泰是头号文豪，他的作品最经得住时间的考验，那么能不能讲一讲他的小说跟微积分有什么关系呢？

师：托尔斯泰(1987)在巨著《战争与和平》中就讲到微积分的应用．他说：人的聪明才智不理解运动(注：相当于光滑曲线)的连续性．人只有在他从某种运动中任意抽出若干单位(注：相当于曲线小段)来进行考察时才逐渐理解．但是把运动分成越来越小的单位，这样处理(注：化为微分)我们只能接近问题的答案，却永远得不到最后的答案．只有采取无限小以及求出它们的总和(注：微分的积分)，我们才能得到问题的答案．数学的一个分支(注：微积分)，已经有了处理无限小求和的技术，从而纠正了人类的智力由于只考察运动的个别单位所不能不犯下的和无法避免的错误．在探讨历史的运动规律时，情况完全一样．由无数人的肆意行为组成的人的运动，是连续的．人的肆意行动的总和永远不能用一个历史人物(一个人、国王或统帅)的活动来表达，只有采取无限小的观察单位——历史的微分，并且运用积分的方法(就是得到这些无限小的总和)，我们才有希望了解历史的规律．这个结论倒像是历史学的基本定理：

历史＝微分的积分，

不能用一个历史人物的活动来表达.

生：但是在前面的“四”中讲的中值定理又说，理论上(在光滑情形)存在某个救世主(只是找不到). 矛盾?

师：这些解读的确牵强附会，但是基本定理毕竟提供一种思维的微元法. 例如，金融危机. 金融数学家彭实戈说：“这是一个非常大的事情，但又是一个非常小的事情. 大楼什么时候倒塌是不确定的事情，确定的是一砖一瓦没有做好，问题出在一砖一瓦没有控制好，才会出现这样的大问题.”

生：一位汽车工业的工程师说，他的计划就是先微分再积分，微分指让每一个工厂做好汽车的某一个原件，积分之后就是一部好汽车. 好像管理学的基本原理也是先做好微分后再积分.

师：也可以说，每人都在做微分，诀窍就在做得最好，精益求精(做成最好的逼近，见前面的“五”中). 所谓细节决定一切，正如微分决定积分.

生：华罗庚讲优选法，讲得最好，讲成世界第一. 一位大学校长讲美化环境，说扫地也要扫成世界第一. 行行出状元!

十一、微积分和自然科学

生：到此，微积分都是取几何图形作参照物，那么微积分只是几何学吗?

师：我们先讲几何，因为图形看得见摸得到，以图代想，看图识字，是认识世界简易可行的方法.

生：那么除了几何应用，还有别的什么应用？

师：微积分更多的应用在于自然科学，它的语言最适合于自然科学，其中斜率或变化率(一阶导函数)，或加速度(二阶导函数)是最基本的概念．自然科学需要变化率、加速度以及微分方程来表示它们的规律，并从中解出它们，见第5章．

生：那么什么是微分方程呢？在中学只知道代数方程和三角方程啊．

师：最简单的微分方程就是已知导函数求函数自身，有最明显的几何意义：已知曲线的斜率求曲线的高，也就是上面“五”中讲的基本定理所做的事．如果退回中学(直角三角形)，一眼认出那就是三角方程：已知斜边的斜率，求它的高．所以大学之中认出有中学．

生：这是我所见过的最形象的对微分方程的解读！

师：每遇到大学新面目，就要认出中学旧面目，新旧对比，温故知新，但是新在哪儿？就要问：

$$大学-中学=?$$

这样才能学到新东西．这也是上面“六”中讨论过的学习方法，大学和中学的关系．

生：那么，除基本定理(已知导函数求函数自身)之外，还有别的微分方程吗？

师：多的是！不过最重要的是人口微分方程（导函数依赖于函数自身，或变化依赖于现状，见G. Strang(1991)著作，它和基本定理一起构成了一阶微分方程的母方程，别的方程都要变成

这两个方程，详见第 5 章．注意，人口方程要用实数，非中学所及！

生：我从电视上听到，微分方程的途径产生了自然科学的革命，由牛顿开始用它，之后，所有的科学家(包括麦克斯韦、薛定谔、爱因斯坦)都追随他，所以牛顿是划时代的．

师：有关多变量的微分方程(包括麦克斯韦、薛定谔和爱因斯坦的微分方程)及其解法乃是数学研究的重大课题．数学太多太难了，这也是华罗庚的感叹．我们的对话只是大海中的一滴水，但愿从一滴能见全身．

生：学海无边！

师：感谢各位[①]参加登山活动以及微积分的问答．

① 他们来自江苏、北京、福建、山东、大连等地（因为尚在试验之中，成败难测，暂不写姓名）．

第2章 微分法(函数体的微分学)

白话不翻译成函数，难以计算.

图形和函数相结合是主流数学的精神. 我们假设同学具备把图形写成函数的能力. 现在能把1.2节由白话翻译成函数.

2.1 相对误差

先定义微分，导数由微分套出.

如图2-1所示，设函数 $f=f(x)$ 定义在闭区间 $[a,b]$ 上（即 $x\in[a,b]$），则在子区间 $[x,x+h]$ 上，对起点 x（固定）及一切附近点 $x+h$（其中 h 是变量）:

$$小段高=f(x+h)-f(x),$$

$$微分=A(x)h,$$

$$测量误差=f(x+h)-f(x)-A(x)h$$

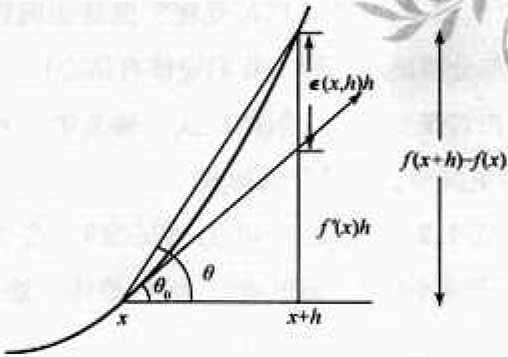


图2-1 相对误差的函数表示

(在 x 的附近点 $x+h$ 处用微分来测量小段高的差量).

相对误差：测量误差与底相比，在 x 及附近点 $x+h$ 处，

$$\epsilon(x,h):=\frac{f(x+h)-f(x)}{h}-A(x), \qquad (2\text{-}1)$$

于是，每一测量误差 =(相对误差)(底)，在 x 及附近点 $x+h$ 处，

$$f(x+h)-f(x)-A(x)h=\epsilon(x,h)h, \qquad (2\text{-}2)$$

其中假设相对误差 $\epsilon(x,h)$ 会小于任一指定的数：

$$|\epsilon(x,h)|\ll 1. \qquad (2\text{-}3)$$

这时，由(2-2)、(2-3)，惟一确定了一个函数

$$f'(x):=A(x),$$

称 $f(x)$ 在点 x 的导函数或斜率(导函数通过微分(2-2)、(2-3)来定义，并不特别，在多变量情形就是如此). 于是，相对误差(2-3)又有了几何解释(1-4)：

$$|\epsilon(x,h)|\leqslant C|\theta-\theta_0|.$$

定义(2-2)、(2-3)是否难以接受呢？但是法国数学课程标准(2001)高三部分都能这么讲，我们应该有信心！

(2-3)是相对误差的最一般的形式，称为第一种. 相对误差还有其他两种更具体、更简单的形式.

第二种：在1.2节之2)看到，在常见情形(连续可微)，相对误差不仅在点点变小(即点态可微)，而且整体一致变小(即一致可微)或

$$|\epsilon(x,h)| \leqslant \epsilon(h) \ll 1 \tag{2-4}$$

(相对误差的上界 $\epsilon(h)$ 跟点 x 无关). 这时微分(2-2)进一步表达为

$$|f(x+h)-f(x)-f'(x)h| \leqslant \epsilon(h)h, \tag{2-5}$$

右端跟点 x 无关.

第三种：相对误差能不能知道是怎么变小的呢？在常见情形(1.11节注6，曲线更光滑)中，不仅有(2-4)，更有可见的相对误差：

$$|\epsilon(x,h)| \leqslant Ch \tag{2-6}$$

与 h 成比例. 这时曲线称为利普希茨可微，(2-5)更明确了：

$$|f(x+h)-f(x)-f'(x)h| \leqslant Ch^2. \tag{2-7}$$

于是，什么叫做 $\ll 1$ 便一眼认出，初学者无需事先学习极限理论.

到此，有三种可微：点态可微(2-2)、(2-3)，一致可微(2-4)、(2-5)和利普希茨可微(2-6)、(2-7). 三种可微由弱到强，由不可见到可见，各有用途. 若考虑充要条件(1-10)，要用(2-3)；

若只考虑充分条件(1-12),便用(2-4);但对于初学者，或者在常见情形下，不妨选用可见误差(2-6)，避免抽象的极限概念.

2.2 导数的推导

试探初等函数，它们的相对误差都如(2-4)或(2-6)，见下面的推导.

采用微分定义(2-2)，使得可能用展开的方法（无须 ε-δ）来构造这个导数：把真高 $f(x+h)-f(x)$ 展开为 h 的一次幂项，其余项是一个高阶的量：比 h 减小得快，那么这个一次幂项的系数(惟一的)便确定为 f 的导数 f'. 什么叫做高阶量或比 h 减小得快呢？举例来说，$f(x)=x^2$，它的真高是 $(x+h)^2-x^2=2xh+h^2$，于是 h 的一次幂项的系数（惟一的）便确定为 $f'(x)=2x$，其余项 $=h^2$，是高阶量(比 h 减小得快)，或者说 $\epsilon(x,h)=h$，跟点 x 无关. 类似地(见2.4节)，对母函数 x^n、$\sin x$、$\cos x$(生成所有的初等函数，暂时不出现 e^x)，都能够按展开法来构造具体的导数 f' 和相对误差 $\epsilon(x,h)$,后者能被 h 的倍数控制：

$$|\epsilon(x,h)|\leqslant Ch,$$

即(2-6). 这就把抽象的 $\epsilon(x,h)$ 换成了具体的 Ch，于是微分(2-2) 换成了(2-7).

注 是否需要极限的问题. 据阿诺德(2006)说：泽利多维奇在自己“为初级物理学家和技术人员”编写的分析学教科书中，把导数确定为函数与自变量的增量比例. 他假设最后的增量

不是特别大. 对于正统数学家们关于需要极限的反驳，他回应道：在这里“比例极限”是没有用的，由于不可能得到太小(比如说,小于 10^{-10}米或秒)的自变量增量，只是因为在这样的尺寸下空间和时间性质成为量子的，所以借助于数学一维连续统描述，超出了模型精确度. 泽利多维奇把“数学导数”理解为方便的近似渐近公式，实际上可以计算我们感兴趣的极限增量的比例.

2.3　微分的变形

现在把 1.3 节前半部由白话翻译成函数.

(2-2) 把微分取在曲线小段的起点 x，只是一个特例，直观可见切线也能取在小段 $[x, x+h]$ 中任一点 $\xi = x + \theta h$ $(0 \leqslant \theta \leqslant 1)$ 处，如图 2-2 所示.

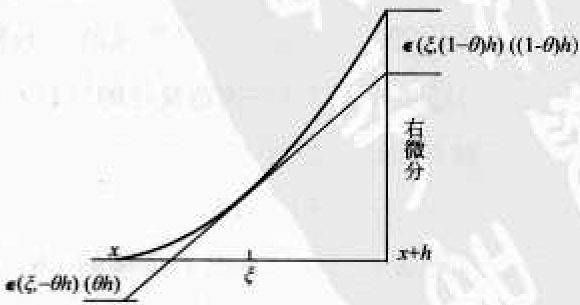


图 2-2　切线取在中间

这时微分(2-2) 能有如下的变形：

$$f(x+h) - f(x) = f'(\xi)h + \bar{\epsilon}(\xi, h)h, \tag{2-8}$$

其中相对误差 ϵ 见(2-9).

证明 只要把 ξ 看成新分点，在两个子区间 $[\xi, x+h]$ 和 $[x,\xi]$ 上分别计算右微分和左微分①

$$f(x+h)-f(\xi)=f'(\xi)((1-\theta)h)+\epsilon(\xi,(1-\theta)h)((1-\theta)h),$$

$$f(\xi)-f(x)=f'(\xi)(\theta h)+\epsilon(\xi,-\theta h)(\theta h),$$

相加即(2-8)，其中相对误差

$$\bar{\epsilon}(\xi,h)=\epsilon(\xi,-\theta h)\theta+\epsilon(\xi,(1-\theta)h)(1-\theta) \tag{2-9}$$

(左右误差的平均). 证毕.

由(2-8)、(2-9)，得知(2-4)~(2-7)有如下变形：

$$|\bar{\epsilon}(\xi,h)|\leqslant\epsilon(h), \tag{2-10}$$

$$|f(x+h)-f(x)-f'(\xi)h|\leqslant\epsilon(h)h, \tag{2-11}$$

$$|\bar{\epsilon}(\xi,h)|\leqslant Ch, \tag{2-12}$$

$$|f(x+h)-f(x)-f'(\xi)h|\leqslant Ch^2. \tag{2-13}$$

它们将在第3章用到.

2.4 导数的更多例子

考察幂函数、三角函数.

1) 当 $f(x)=x^n(n=0, 1, 2, \cdots)$，它的真高是

$$(x+h)^n-x^n=nx^{n-1}h+\frac{n(n+1)}{2!}x^{n-2}h^2+\cdots+h^n,$$

① 先有右微分：$f(x+h)-f(x)-f'(x)h=\epsilon(x,h)h$，再有左微分(把 h 改成 $-h$)：$f(x)-f(x-h)-f'(x)h=\epsilon(x,-h)h$.

则一次幂项的系数(惟一的)便是 $f'(x)=nx^{n-1}$，其余项

$$\left|\frac{n(n+1)}{2!}x^{n-2}h^2+\cdots+h^n\right|\leqslant Ch^2$$

是高阶量(比 h 减小得快). 或者说测量误差 $|\epsilon(x,h)|\leqslant Ch$.

2）当 $f(x)=\sin x$，则选(惟一的) $f'(x)=\cos x$，是高阶量(比 h 减小得快). 暂设 $h>0$（自己检查 $h<0$ 的情形），

$$\begin{aligned}&|\sin(x+h)-\sin x-f'(x)h|\\&=|\sin x\cos h+\cos x\sin h-\sin x-\cos x\cdot h|\\&=|\sin x(\cos h-1)+\cos x(\sin h-h)|\\&\leqslant 1-\cos h+(h-\sin h)\leqslant Ch^2,\end{aligned}$$

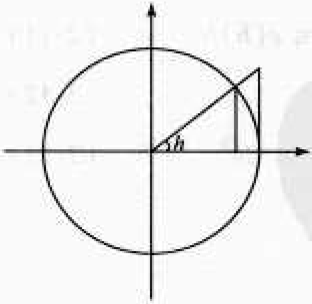


图 2-3 三角不等式

$\sin h<h<\tan h$

这里用到三角不等式(图 2-3)：

$$1-\cos h=2\sin^2\frac{h}{2}\leqslant\frac{1}{2}h^2,$$

$$h-\sin h<h-h\cos h\leqslant\frac{1}{2}h^3,$$

或

$$\epsilon(x,h)=\frac{1-\cos h}{h}+\frac{h-\sin h}{h}\leqslant Ch.$$

类似地(作为练习)，当 $f(x)=\cos x$，可选(惟一的) $f'(x)=-\sin x$，使测量误差 $\leqslant Ch^2$，是高阶量（比 h 减小得快），或者说相对误差 $\epsilon(x,h)\leqslant Ch$.

三角函数求导的更有趣的方法，见 M. Livshits(2004).

但是，对于对数函数 $\ln x$ 和指数函数 e^x，其定义到 3.5 节再给出.

下面列出最简单的一份导数表：

$$(x^n)' = nx^{n-1},$$
$$(\sin x)' = \cos x,$$
$$(\cos x)' = -\sin x,$$

其中，第一式中的 n 不仅是自然数，也可以是有理数(其中 $x>0$，见2.5节之4))，甚至是无理数.(但其定义要用到实数,另册再讲).都怪这些公式不够美(例如,第二、第三式差一符号),经久不用难免记不清. 但是有一条，e^x 的导数不变，不会记错. 这是忘了其他函数的导数后还能留下的惟一结果.

易见 sin 和 cos 都满足下述的“二阶微分方程”：

$$f''(x) = -f(x).$$

它们的线性组合显然也满足，反之亦真，见5.2节.

2.5　微分法则

原则上，可以继续用展开法对更多的具体函数构造导数. 如果这样，对于每一个具体函数，都要有一种方法来构造它的导数和相对误差，太费事了，没完没了. 微分法则之所以成功，正是因为它只依赖于非常少的几个初等函数，如多项式、正弦函数和3.5节的指数函数(称为母函数),不用每次都从定义出发来求函数的导数，而是通过下述微分法则，生成任何初等函数的导数.

函数加减乘除的导数、复合函数以及反函数的导数，都有一定的法则来算.

(i) 加法法则 $(f+g)' = f' + g'$,数乘法则 $(cf)' = cf'$. 这两

个法则好想像到；

（ii）乘法法则$(fg)'=f'g+fg'$，它不是希望的那样：$(fg)'=f'g'$；

（iii）除法法则：若$|f(x)|\geqslant m>0$，则

$$\left(\frac{1}{f}\right)'=-\frac{f'}{f^2},\quad \left(\frac{g}{f}\right)'=\frac{g'f-gf'}{f^2},$$

它不是希望的那样：$\left(\frac{g}{f}\right)'=\frac{g'}{f'}$.

多次应用法则(ii)就能得出求幂法则：

$$((f(x))^n)'=n(f(x))^{n-1}f'.$$

但有更一般的复合函数的导数：

（iv）链式法则：

$$(f(g(x)))'=f'(g(x))g'(x).$$

这是最难的微分法则，但不可少，因为大多数新函数的构造来源于链式函数$f(g(x))$，即把一个函数插入另一个函数之中，称为复合. 首先要弄清楚函数之间的组合. 例如，

$$y=\sin x^2\quad 重写作\quad y=\sin z,\quad z=x^2,$$

$$y=\sin^2 x\quad 重写作\quad y=z^2,\quad z=\sin x,$$

$$y=\sqrt{1-x^2}\quad 重写作\quad y=z^{1/2},\quad z=1-x^2.$$

（v）反函数求导：$g(x)$是f的反函数指f和g相互抵消：$f(g(x))=x$且$g(f(x))=x$，由后者两边求导，并用(iv)，则当$|f'(x)|\geqslant m>0$，有$g'(x)=1/f'(g(x))$.

这些法则很实用，推导并不难，只是机械的演算（见本章附录），但演算较长，有点乏味，只要知道确有方法能够生成任何

初等函数(多项式、有理函数、三角函数、指数函数)的导数就可以了．例如，

1）从$(x)'=1$和乘法法则(ii)，利用数学归纳(作为练习)能够得到x^n可微，并有$(x^n)'=nx^{n-1}$.

2）多项式函数可微，并有

$$\begin{aligned}&(a_0+a_1x+a_2x^2+\cdots+a_nx^n)'\\&=a_1+2a_2x+3a_3x^2+\cdots+na_nx^{n-1}.\end{aligned}$$

3）从$(x)'=1$和除法法则（iii)，利用数学归纳(作为练习)能够得到$1/x^n$在任何不包含0的闭区间可微，并有

$$\left(\frac{1}{x^n}\right)'=\frac{-n}{x^{n+1}}\quad 或\quad (x^{-n})'=-nx^{-n-1},$$

美妙的是这与1）具有相同的形式：乘以指数并且指数减一.

所以，幂函数的微分法则$(x^n)'=nx^{n-1}$，无论n是正还是负都成立，但n必须是整数.

4）我们希望指数为分数$n=p/q$时，能保持相同的法则：$x^n=nx^{n-1}$. 令$f=x^{p/q}$，即$f^q=x^p$：

$$qf^{q-1}f'=px^{p-1}\quad（两边运用求幂法则），$$

$$f'=\frac{px^{-1}}{qf^{-1}}\quad（约分，注意x^p=f^q），$$

$$f'=nx^{n-1}\quad（注意n=\frac{p}{q},\quad f=x^n），$$

所以，幂函数的微分法则$(x^n)'=nx^{n-1}$，当n是分数时也成立(实际上，n可以是任意实数).

5）再利用加法(i)和数乘法则(ii)以及除法法则(iii)能够得

到有理函数(即两个多项式之比)在分母不为0的任意闭区间上的微分法则. 证明作为练习.

6) 2.4节已经知道 $\sin x$ 和 $\cos x$ 的导数. 利用除法法则(iii)还能够得到 $\tan x$、$\cot x$、$\sec x$、$\csc x$ 的导数:

$$(\tan x)' = \sec^2 x,$$

$$(\cot x)' = -\csc^2 x,$$

$$(\sec x)' = \sec x\tan x,$$

$$(\csc x)' = -\csc x\cot x.$$

证明作为练习. 记不住就查表.

7) 如果 $y = \sin x, x = \arcsin y$,那么

$$\frac{dy}{dx} = y' = \cos x,$$

$$\frac{dx}{dy} = \frac{1}{\cos x} = \frac{1}{\sqrt{1-\sin^2 x}} = \frac{1}{\sqrt{1-y^2}}.$$

这里还引用以下6个函数作为例子(细节见G. Strang(1991)).

函数 $f(y)$	导数 dx/dy	函数 $f(y)$	导数 dx/dy
$\arcsin y$	$1/\sqrt{1-y^2}$	$\text{arccot}\, y$	$-1/(1+y^2)$
$\arccos y$	$-1/\sqrt{1-y^2}$	$\text{arcsec}\, y$	$1/(\lvert y\rvert\sqrt{y^2-1})$
$\arctan y$	$1/(1+y^2)$	$\text{arccsc}\, y$	$-1/(\lvert y\rvert\sqrt{y^2-1})$

证明作为练习. 记不住就查表.

这些导数都是在微分学中经常用到的.

导数的应用(如极大和极小理论)在第1章附录中有些描述,但主要参考其他教材,例如,国内有华罗庚(1963年以来)的《高

等数学引论》，国外有：G. Strang（1991）的 *Calculus*，Y. Zeldovich - I. Yaglom（1987）的 *Higher Math for Beginners*. 教师可结合不同专业（如物理、生命科学、经济、管理）来选取.

在结束本章时，我们强调一下，从那些具体函数构造导数的过程中，认可了微分各种形式（2-2）、（2-5）、（2-7），也包括（2-8）、（2-11）、（2-13）的可行性.

附录　微分法则的推导

为了节省符号和计算，我们采用如下形式的利普希茨可微的定义(2-7)：

$$f(x+h)-f(x)-f'(x)h=O(h^2),$$

$$|O(h^2)|\leqslant Ch^2.$$

加法和数乘法则(i)易证(作为练习). 现证乘法法则(ii)：

$$\begin{aligned}
&f(x+h)g(x+h)-f(x)g(x)\\
&=f(x+h)g(x+h)-f(x+h)g(x)+f(x+h)g(x)-f(x)g(x)\\
&=f(x+h)(g(x+h)-g(x))+g(x)(f(x+h)-f(x))\\
&=f(x)(g(x+h)-g(x))+g(x)(f(x+h)-f(x))\\
&\quad+(f(x+h)-f(x))(g(x+h)-g(x))\\
&=f(x)(g'(x)h+O(h^2))+g(x)(f'(x)h+O(h^2))\\
&\quad+(f'(x)h+O(h^2))(g'(x)h+O(h^2))\\
&=(f(x)g'(x)+f'(x)g(x))h+O(h^2).
\end{aligned}$$

所以，fg 可微，并有 $(fg)'=f'g+fg'$.

除法法则（iii）的证明：能从$(f)(1/f)=1$开始，注意1的导数是0，再应用乘法法则：

$$f\left(\frac{1}{f}\right)'+\frac{1}{f}f'=0 \quad \text{所以} \quad \left(\frac{1}{f}\right)'=\frac{-f'}{f^2}.$$

也能由导数的原始定义开始：

$$\frac{1}{f(x+h)}-\frac{1}{f(x)}$$

$$=\frac{f(x)-f(x+h)}{f(x+h)f(x)}$$

$$=\frac{-f'(x)h+O(h^2)}{f(x+h)f(x)}$$

$$=\frac{-f'(x)h+O(h^2)}{(f(x))^2}+\frac{-f'(x)h+O(h^2)}{f(x+h)f(x)}$$

$$-\frac{-f'(x)h+O(h^2)}{(f(x))^2}$$

$$=\left(-\frac{f'(x)}{(f(x))^2}\right)h+\frac{O(h^2)}{(f(x))^2}$$

$$+\left(\frac{-f'(x)h+O(h^2)}{f(x)}\right)\left(\frac{1}{f(x+h)}-\frac{1}{f(x)}\right)$$

$$=\left(-\frac{f'(x)}{(f(x))^2}\right)h+O(h^2)+O(h)\frac{f(x)-f(x+h)}{f(x+h)f(x)}$$

$$=\left(-\frac{f'(x)}{(f(x))^2}\right)h+O(h^2).$$

所以，$1/f$可微，并有$(1/f)'=-f'/f^2$.

再联合乘法法则（ii）能证得（作为练习）一般的除法法则：

$$\left(\frac{g}{f}\right)'=\frac{g'f-gf'}{f^2}.$$

求幂法则是乘法法则(ii)的推论. 最好的途径是采用数学归纳法. 当 $k=1$ 时,简化为 $f'=f'$. 当 $k=2$ 时,得到 $(f^2)'=2ff'$. 假设 $k=n$ 时求幂法则成立,由乘法法则推出 $k=n+1$ 时也成立:

$$(f^{n+1})'=(f^nf)'=f^nf'+f(nf^{n-1}f')=(n+1)f^nf'.$$

这样就得到正整数的求幂法则. 类似的, 从指数 -1 开始, 我们又能从除法法则得到负整数的求幂法则. 当指数为分数 $n=p/q$ 时, 令 $g=f^{p/q}$,即 $g^q=f^p$:

$$qg^{q-1}g'=pf^{p-1}\quad\text{(两边运用整数的求幂法则)},$$

$$g'=\frac{pf^{-1}}{qg^{-1}}\quad\text{(约分,注意 } f^p=g^q\text{)},$$

$$g'=nf^{n-1}\quad\left(\text{注意 } n=\frac{p}{q},\ \ g=f^n\right).$$

所以, 求幂法则 $(f^n)'=nf^{n-1}$, 当 n 是分数时也成立 (实际上, n 能是任意实数);

链式法则 (iv) 的证明:

$$\begin{aligned}g(f(x+h))-g(f(x)) &= g'(f(x))(f(x+h)-f(x))+O(h^2)\\ &= g'(f(x))(f'(x)h+O(h^2))+O(h^2)\\ &= g'(f(x))f'(x)h+g'(f(x))O(h^2)+O(h^2)\\ &= g'(f(x))f'(x)h+O(h^2).\end{aligned}$$

反函数微分法则(v)的证明: 如果 $f(g(x))=x$ 且 $g(f(x))=x$,那么 $g(x)$ 就是 $f(x)$ 的反函数. 如果 $f(x)$ 可微且导数满足 $f'\geq m>0$,由基本定理知道 f 严格递增. 因此 f 有反函数 $g(x)$ 在 $[f(a),f(b)]$ 上. 利用 $f(g(x))=x$ 和链式法则 $f'(g(x))g'(x)=1$,证得 $g(x)$ 可微且有

$$g'(x)=\frac{1}{f'(g(x))}\quad (f(g(x))=x).$$

令 $y=g(x)$ 和 $x=f(y)$，能够写成更好看的形式：

$$\frac{dx}{dy}\frac{dy}{dx}=1 \quad 或 \quad \frac{dx}{dy}=\frac{1}{dy/dx},$$

即 $x=g^{-1}(y)$ 的斜率乘以 $y=g(x)$ 的斜率等于 1.

图解的证明见 M. Livshits（2004）.

第 3 章
积分法(函数体的积分学)

白话不翻译成函数，教师不放心.

第 2 章只考虑小段区间上的高以及微分. 本章要考虑全段区间上的

$$全段高 = f(b) - f(a)$$

以及各小段上

$$微分的和 = \sum_{x \in 各分点} f'(x)h,$$

这里用记号 $\sum$ 来缩短加号的写法. 全段高跟微分和之间有全段误差:

$$f(b) - f(a) - \sum_{x \in 各分点} f'(x)h.$$

3.1　基本定理: 简单形式

现在把 1.2 节由白话翻译成函数.

计算全段误差，关键技术是转换恒等式(1-8)或

$$f(b)-f(a)-\sum_{x\in\text{各分点}} f'(x)h$$

$$=(b-a)(\epsilon(x,h)\text{ 的平均}), \quad (3\text{-}1)$$

转换为第2章有过的相对误差. 当且仅当有可积条件(1-9)(这样的曲线称为平均可微)：对所有分割当加密

$$|\epsilon(x,h)\text{ 的平均}| \leqslant \epsilon(h) \ll 1 \quad (3\text{-}2)$$

(平均的上界与分点 x 无关) 才有基本定理.

如果加强些(不是充要条件了)，直接用上界(2-4)：

$$|\epsilon(x,h)| \leqslant \epsilon(h) \ll 1$$

(相对误差的上界与分点 x 无关)，于是平均$\leqslant\epsilon(h)$，(1-13)化为

$$\left|f(b)-f(a)-\sum_{x\in\text{各分点}} f'(x)h\right| \leqslant (b-a)\epsilon(h) \quad (3\text{-}3)$$

与相对误差的上界成正比.

这个上界 $\epsilon(h)$ 有多小? 直到4.1节才认出

$$\epsilon(h)=\underset{s,t\in[x,x+h],x\in[a,b]}{\text{upper}} |f'(s)-f'(t)|.$$

但是，如果有利普希茨条件(2-6)，(3-2)变成显示的估计：

$$\left|f(b)-f(a)-\sum_{x\in\text{各分点}} f'(x)h\right| \leqslant Ch \quad (3\text{-}4)$$

与最大底成正比.

副产品得到微分学中两个最有用的定理：

$$f'(x)\equiv 0 \Rightarrow f(x)=c,$$

$$f'(x)>0 \Rightarrow f(x)\uparrow$$

(当 f' 可积).

3.2 基本定理：完整形式

现在把1.3节的后半部由白话翻译成函数.

由微分的变形(2-8)出发，那么全段高跟完整的微分和(即现有教材的黎曼和)之间有全段误差

$$\begin{aligned} & f(b)-f(a)-\sum_{\xi\in\text{中间点}} f'(\xi)h \\ &= \sum_{\xi\in\text{中间点}} h\bar{\epsilon}(\xi,h) \\ &= (b-a)(\bar{\epsilon}\text{的平均}). \qquad (3\text{-}5) \end{aligned}$$

这实际是(3-1)的变形：由特殊和变到黎曼和，即是变回了相对误差$\bar{\epsilon}$，取其平均. 当且仅当有(1-9)(曲线平均可微)，才有基本定理. 所以，把微分定义(2-2)变形为(2-8)，才有完整的基本定理!

由于(3-5)太重要了，下面仔细写下全部证明过程：

把全区间$[a,b]$分割成$n+1$小段区间$[x_i,x_i+h_i]$，其中$a=x_0<x_1<\cdots<x_{n+1}=b$，$h_i=x_{i+1}-x_i$(图3-1).



图3-1 区间分割

在各小段区间上写出小段高的测量误差(2-8)，即是

$$f(x_0+h_0)-f(x_0)-f'(\xi_0)h_0=\bar{\epsilon}(\xi_0,h_0)h_0,$$

$$f(x_1+h_1)-f(x_1)-f'(\xi_1)h_1=\bar{\epsilon}(\xi_1,h_1)h_1,$$

$$\cdots = \cdots$$

$$f(x_n + h_n) - f(x_n) - f'(\xi_n)h_n = \bar{\epsilon}(\xi_n, h_n)h_n.$$

然后把它们一个挨一个地加起来，f 的中间值相消，剩下两头，即全区间上的总高 $f(b)-f(a)$. 其结果，左边写成

$$f(b) - f(a) - \sum_{i=0}^{n} f'(\xi_i)h_i.$$

右边相加，便是全段误差（3-5）：

$$\sum_{i=0}^{n} \bar{\epsilon}(\xi_i, h_i)h_i = (b-a)\sum_{i=0}^{n} \frac{h_i}{b-a}\bar{\epsilon}(\xi_i, h_i).$$

此即相对误差 $\bar{\epsilon}$ 的平均（差一倍数 $b-a$）. 当且仅当有可积条件（1-9）（曲线平均可微）：对所有分割当加密

$$|\bar{\epsilon}(\xi_i, h_i)\text{ 的平均}| \leqslant \epsilon(h) \ll 1 \qquad (3\text{-}6)$$

（平均的上界与点 ξ 无关）才有基本定理. 这里和以后把 h 理解为最大的 h_i.

如果加强些（不是充要条件了），直接用上界（2-10）（即一致可微）：

$$|\bar{\epsilon}(\xi_i, h_i)| \leqslant \epsilon(h), \quad i = 0, \cdots, n$$

（相对误差的上界与点 ξ 无关），于是平均 $\leqslant \epsilon(h)$，（1-11）化为

$$\left| f(b) - f(a) - \sum_{i=0}^{n} f'(\xi_i)h_i \right| \leqslant (b-a)\epsilon(h),$$

与相对误差的上界成正比. 如果有利普希茨条件（2-10），更具体化为

$$\left| f(b) - f(a) - \sum_{i=0}^{n} f'(\xi_i)h_i \right| \leqslant Ch,$$

与最大底成正比.

当不断细分 h_i，右端便可略去不计，便有理想的等式

$$f(b)-f(a)=\int_a^b f'(x)\,\mathrm{d}x \tag{3-7}$$

(数学是一门理想的学问)，这里把离散的和写为连续的积分，它是一个确定的值，不依赖于如何细分以及如何取中间值(不一定取在分点 x_i)，总是趋向一个确定的数 $f(b)-f(a)$.

后记[①] 如果曾学过概率，上述基本定理的可积条件(3-6)(其中分点随机选取)可以用概率来解释．如果随机变量 x 的概率密度为

$$p(x)=\begin{cases}1, & 0\leqslant x\leqslant 1\\ 0, & 其他,\end{cases}$$

则有以下结论：任意随机函数 $f(x)$ 的期望为 $E(f(x))=\int_{-\infty}^{\infty} f(x)p(x)\,\mathrm{d}x=\int_0^1 f(x)\,\mathrm{d}x$．由上述，分割是任意的，分点位置 x 是具有上述概率密度的随机变量．所以，当 h 给定时，相对误差 $\epsilon(x,h)$ 就是关于 x 的随机函数，它的期望为 $E(\epsilon(x,h))=\int_0^1 \epsilon(x,h)\,\mathrm{d}x$．若 $\epsilon(x,h)$ 与 x 无关，便能将其上界 $\epsilon(h)$ 代入上式右端，得到 $E(\epsilon(x,h))=\int_0^1 \epsilon(x,h)\,\mathrm{d}x\leqslant\epsilon(h)$．对任意的区间 $[a,b]$，只需在上式右端乘上区间长度(常数)即可．这样，当 $h\to 0$ 时，$\epsilon(h)\to 0$，所以期望 $E(\epsilon(x,h))$，即全段误差也趋于0，基本定理成立.

① 这个后记是由朱起定改写的.

3.3 基本定理：充分条件(初学不读)

1) 3.2 节已给出了最简单的充分条件(2-4)，即 f 一致可微. 这不需要实数或其他知识.

下面利用 M. Livshits(2004)的巧妙方法（避开中值定理，但要用一点实数），证明其他几个充分条件. M-Livshits 先证明了下面的增函数定理：

假设导数 $f'(x)\geqslant 0, x\in[a,b]$，那么 f 是增函数：$f(b)\geqslant f(a)$.

上面定理貌似平凡(图形易知)，但证明不凡. 华罗庚(1963，第六章)从 $f'(a)$ 的定义出发：

$$f(a+h)=f(a)+f'(a)h+\epsilon(a,h)h.$$

如果 $f'(a)>0$，取 $-\epsilon(a,h)<f'(a)$（当 h 充分小），所以 $f(a+h)>f(a)$. 这是不是说明 f 在 a 附近上升呢？数学需要严格的拷问. 张景中也指出，它并不说明在区间 $(a,a+h)$ 中任取两点 $u<v$ 都有 $f(u)<f(v)$（除非 f 一致可微）. 所以增函数定理还需要更严格的证明. 其证明有点技巧，初学先不读.

先设 $f'\geqslant c>0$ 在 $[a,b]$，引入

$$S=\{x\in[a,b]: f(x)\geqslant f(a)\}.$$

既然 $a\in S$，那么 $S\neq\varnothing$. 令 $d=\sup S$，则有 $f(d)\geqslant f(a)$（由于 f 连续）. 如果 $d<b$，从上述点态可微的定义得到

$$f(d+h)\geqslant f(d)+f'(d)h-\epsilon(d,h)h\geqslant f(d)+(c-\epsilon(d,h))h,$$

所以可以找到某个 $h>0$ 使得

$$f(d+h)\geqslant f(d)\geqslant f(a),$$

所以 $d+h\in S$. 这与 $d=\sup S$ 矛盾！于是 $d=b$，即 $f(b)\geqslant f(a)$.

证明有点形式，意思是要证 $f(x)\geqslant f(a)$ 对一切 $x\in[a,b]$，只要先证使之成立的区间 $[a,y]$，y 的上确界为 b.

如果只有 $f'\geqslant 0$，引入辅助函数 $g(x)=f(x)+cx,c>0$. 此时回到上述情形，最后得到

$$f(b)-f(a)\geqslant c(a-b).$$

既然 c 是任意常数，$f(b)\geqslant f(a)$.

还可以得到增函数定理的等价形式

$$f'(t)\geqslant g'(t),\ t\in[a,b]\Rightarrow f(b)-f(a)\geqslant g(b)-g(a).$$

只需将上述证明左端换成 $f-g$ 即得.

然后，M. Livshits(2004) 再来证明：

2）f' 连续等价于 f 一致可微.

必要性很显然，只要将 (2-8) 中的 ξ 换成 $\eta(\xi,\eta\in[x,x+h])$：

$$\left|\frac{f(x+h)-f(x)}{h}-f'(\eta)\right|\leqslant\epsilon(h).$$

由三角不等式便得

$$|f'(\xi)-f'(\eta)|\leqslant\epsilon(h) \tag{3-8}$$

（但差一个倍数）.

充分性要有一点小技巧. 假设 f' 一致连续：

$$f'(\xi)-\epsilon(h)\leqslant f'(t)\leqslant f'(\xi)+\epsilon(h),\quad t,\xi\in[x,x+h].$$

注意到 ξ 和 h 与 t 无关，于是 $f'(\xi)$ 和 $\epsilon(h)$ 都是常数，应用增（或减）函数定理（其中的 $g(t)=(f'(\xi)\pm\epsilon(h))t$）得到

$$(f'(\xi)-\epsilon(h))h \leqslant f(x+h)-f(x) \leqslant (f'(\xi)+\epsilon(h))h.$$

于是

$$|\bar{\epsilon}(\xi,h)| \leqslant \epsilon(h),$$

所以f一致可微，且有一致估计

$$\left|f(b)-f(a)-\sum_{i=0}^{n} f'(\xi_i)h_i\right| \leqslant (b-a)\epsilon(h). \qquad (3\text{-}9)$$

3）f'只有有限个间断.

当f'有界，并且除去有限个间断外，在其他各点连续，基本定理同样成立. 证明如下(图 3-2)：



图 3-2 区间分点

把$[a,b]$分成$n+1$个小段区间，分点记作$x_0=a, x_1, \cdots, x_n, x_{n+1}=b$,其中$x_i \leqslant \xi_i \leqslant x_{i+1}$, $h_i=x_{i+1}-x_i$, $h=\max h_i$, $i=0,1,\cdots,n$. 设$|f'(x)| \leqslant C$, $x \in [a,b]$,且f'的间断点只有一个(多于一个的情况可以类似得到),记作p，可知与点p相交的区间至多有 2 个，如图 3-3、图 3-4 所示. 区间总和记作J，总长度不超过$2h$.

当间断点p恰好是分点时，J包含两个区间.

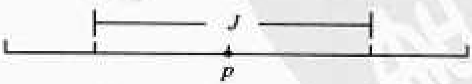


图 3-3 间断点是分点

当间断点p不是分点时，J只包含一个区间.

在$[a,b]$减去J余下的有限个闭子区间上，函数f'连续，所



图3-4　间断点不是分点

以 f 一致可微，相对误差一致小，即当 $\xi_i\in[a,b]/J$，

$$|\bar{\epsilon}(\xi_i,h_i)|\leqslant\epsilon(h)\ll 1. \tag{3-10}$$

在 J 上(f'有界)可以证明相对误差有界.

由 $|f'|\leqslant C$ 和三角不等式得到

$$f'(\xi_i)-2C\leqslant f'(t)\leqslant f'(\xi_i)+2C\quad t,\ \xi_i\in[x_i,x_{i+1}].$$

注意到 ξ_i 与 t 无关，于是 $f'(\xi_i)$ 是常数，应用增(或减)函数定理(其中的 $g(t)=(f'(\xi_i)\pm 2C)t$)得到

$$(f'(\xi_i)-2C)h_i\leqslant f(x_{i+1})-f(x_i)\leqslant(f'(\xi_i)+2C)h_i,$$

所以

$$|\bar{\epsilon}(\xi_i,h_i)|\leqslant 2C. \tag{3-11}$$

综合两种情况，平均（或全段误差）能变小：

$$\begin{aligned}\left|\sum\bar{\epsilon}(\xi_i,h_i)h_i\right| &\leqslant\sum{}_{(1)}|\bar{\epsilon}(\xi_i,h_i)|h_i+\sum{}_{(2)}|\bar{\epsilon}(\xi_i,h_i)|h_i\\ &\leqslant\epsilon(h)\sum{}_{(1)}h_i+2C\sum{}_{(2)}h_i\\ &\leqslant\epsilon(h)(b-a)+2C\cdot 2h\ll 1,\end{aligned} \tag{3-12}$$

其中第一部分 $\sum_{(1)}$ 所涉及的小区间不与间断点 p 相交，有(3-10)；第二部分 $\sum_{(2)}$ 所涉及的每一小区间属于 J，都与间断点 p 相交，有（3-11）. 便可以得出估计式

$$\left|f(b)-f(a)-\sum_{i=0}^{n}f'(\xi_i)h_i\right|\leqslant(b-a)\epsilon(h)+C\cdot h. \tag{3-13}$$

4）f' 单调（可有无限个间断）.

最有趣的条件是f'单调有界，不妨认为单调不减. 把$[a,b]$分成$n+1$个小段区间，分点记作$x_0=a,x_1,\cdots,x_n,x_{n+1}=b$,其中$x_i\leqslant\xi_i\leqslant x_{i+1}$, $h_i=x_{i+1}-x_i$,$h=\max h_i$, $i=0,1,\cdots,n$. 如图 3-5 所示.

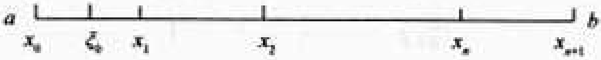


图 3-5　区间分点

由f'单调不减，得

$$f'(x_i)\leqslant f'(t)\leqslant f'(x_{i+1}),\quad t\in[x_i,x_{i+1}]. \tag{3-14}$$

注意到x_i, x_{i+1}与t无关，于是$f'(x_i)$和$f'(x_{i+1})$都是常数，应用增(或减)函数定理（其中的$g(t)=f'(x_{i+1})t$或$g(t)=f'(x_i)t$）得到

$$f'(x_i)h_i\leqslant f(x_{i+1})-f(x_i)\leqslant f'(x_{i+1})h_i,$$

便有

$$f'(x_i)-f'(\xi_i)\leqslant\bar{\epsilon}(\xi_i,h_i)\leqslant f'(x_{i+1})-f'(\xi_i),$$

所以有

$$\begin{aligned}|\bar{\epsilon}(\xi_i,h_i)|&\leqslant\max(|f'(x_{i+1})-f'(\xi_i)|,|f'(x_i)-f'(\xi_i)|)\\&\leqslant f'(x_{i+1})-f'(x_i)\end{aligned}\tag{3-15}$$

未必小，但是有界，而且其和也有界：

$$\begin{aligned}\left|\sum\bar{\epsilon}(\xi_i,h_i)\right|&\leqslant\sum[f'(x_{i+1})-f'(x_i)]\\&\leqslant f'(b)-f'(a).\end{aligned}\tag{3-16}$$

所以经平均(乘上h_i)，全段误差很小：

$$\left|f(b)-f(a)-\sum_{i=0}^{n}f'(\xi_i)h_i\right| \leqslant |f'(b)-f'(a)|h. \tag{3-17}$$

用可见的量躲开抽象的 $\varepsilon-\delta$ 语言. 以上对3)、4)的证明是由我的学生陈竑焘做的.

这一节的证明有点长，幸亏结论简明.

3.4 由微分表得积分表

(3-7)给出了积分法：如果积分号下的函数可以写成导函数 $f'(x)$ 的形式，那么积分的结果便是原来函数 $f(x)$ 在两个端点值的差 $f(b)-f(a)$. 这样，把2.4节的微分表倒过来，就得到了积分表. 下面偷懒，在（3-7）中略去积分的上下限及左边原来函数的差值，直接写成 $\int f'(x)\mathrm{d}x = f(x)$：

$$\int x^n \mathrm{d}x = \frac{x^{n+1}}{n+1}, \quad n \neq -1,$$

$$\int \cos x \mathrm{d}x = \sin x,$$

$$\int \sin x \mathrm{d}x = -\cos x$$

(注意：这不是现有教材中的不定积分——本书不需要这个概念).

3.5 积分表的突破：自然对数和指数（有赖于实数）

诚如 M. Livshits(2004)指出的，3.4节的第一个积分，虽然

当 $n=-1$ 时失去意义，但积分号下的函数 $1/t(1\leqslant t\leqslant x)$ 是一个好函数,应该有积分(或面积)，记之为自然对数函数

$$\ln(x)=\int_1^x\frac{1}{t}\mathrm{d}t. \tag{3-18}$$

虽然直观上深信，这个函数的面积应该存在，但我们还是要通过代数，计算它的左(端点)和与右(端点)和(图 3-6、图 3-7)：

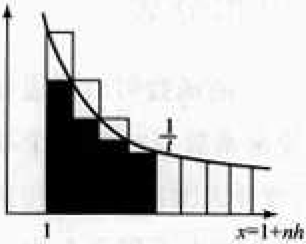


图 3-6　左和与右和

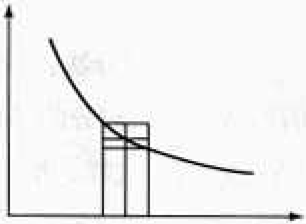


图 3-7　左右和的单调性

左和为

$$L_n=h\left(1+\frac{1}{1+h}+\frac{1}{1+2h}+\cdots+\frac{1}{1+(n-1)h}\right),$$

右和为

$$R_n=h\left(\frac{1}{1+h}+\frac{1}{1+2h}+\cdots+\frac{1}{1+(n-1)h}+\frac{1}{1+nh}\right).$$

显然

$$R_n<L_n,\quad L_n-R_n=h\left(1-\frac{1}{x}\right).$$

由图 3-7 还能看出左右和的单调性：

$$R_n<R_{2n}<L_{2n}<L_n.$$

代数的证明基于一般项

$$\frac{h}{2}\left[\frac{1}{1+(2m)\cdot h/2}+\frac{1}{1+(2m+1)\cdot h/2}\right]<\frac{h}{1+mh},\quad 0\leqslant m\leqslant n-1,$$

$$\frac{h}{1+mh}<\frac{h}{2}\left[\frac{1}{1+(2m-1)\cdot h/2}+\frac{1}{1+(2m)\cdot h/2}\right],\quad 1\leqslant m\leqslant n,$$

所以，只要加入新分点，右和增加，左和减少，它们彼此靠近，但并不越过：

$$R_n<R_{2n}\leqslant \qquad \leqslant L_{2n}<L_n$$

中间夹住的一个数(根据实数的完备性,或说实轴无"洞")猜想就是 $1/t(1\leqslant t\leqslant x)$ 的积分,或对数函数(3-18)(对于一般的连续函数,也有这样的结果,见 G. Strang(1991),但我们针对特殊的 $1/t$ 来验证，更简单了).

G. Strang(1991)提出一个问题：这个数与加密方式有关吗?回答是否定的. 事实上，对于一般的加密方式，由图3-6、图3-7，同样得到左右和序列，并有性质：只要加入新分点，右和增加，左和减少，它们彼此靠近，但并不越过①，中间也夹住一个数.下面用反证法，证明对于任一加密方式，夹住的数都是惟一的.

反证法：不妨把上述通过等距对半加密方式夹住的数记作 A. 假设存在另一种加密方式(*),得到的左右和序列分别是

---

① 证明如下:因为对于 $1\leqslant\alpha,\beta\leqslant x$,有

$$\left|\frac{1}{\alpha}-\frac{1}{\beta}\right|\leqslant\frac{|\alpha-\beta|}{\alpha\beta}\leqslant|\alpha-\beta|,$$

所以对于任意的加密方式,其中任一子区间 $[t_i,t_j]$ 的左右和相差 $(1/t_i-1/t_j)(t_j-t_i)\leqslant(t_j-t_i)^2$,因而整个区间 $[1,x]$ 的左右和相差

$$\sum\left(\frac{1}{t_i}-\frac{1}{t_j}\right)(t_j-t_i)\leqslant\sum(t_j-t_i)^2\leqslant(x-1)\max|t_j-t_i|,$$

所以当 $\max|t_j-t_i|$ 减小时,左右和序列彼此靠近,但并不越过.

$\{L_n^*, R_n^*\}$，夹住的数是 $B(B\neq A)$；那么每一次都将这两种加密方式(等距对半加密方式和另一种加密方式(∗))的分点合并，就得到一种新的加密方式(†)，得到的左右和序列记作 $\{R_n^\dagger, L_n^\dagger\}$. 显然它既是对等距对半加密方式的加密，也是对加密方式(∗)的加密，便有

$$R_n\leqslant R_n^\dagger\leqslant L_n^\dagger\leqslant L_n,$$
$$R_n^*\leqslant R_n^\dagger\leqslant L_n^\dagger\leqslant L_n^*.$$

由于 $\{R_n, L_n\}$ 夹住数 A，$\{R_n^*, L_n^*\}$ 夹住数 B，从而对于加密方式(†)，左右和序列 $\{R_n^\dagger, L_n^\dagger\}$ 夹住两个不同的数 A 和 B，与任意加密方式的左右和序列彼此靠近矛盾！这就证明不论通过何种加密方式，得到的左右和序列夹住的数(即积分)是惟一的.

以上证明有点啰嗦，但中间并无跨度，是自然过渡，只需要耐心！

对数的如此定义虽然是构造性的，但并非实用算法. 待到第4章泰勒级数时，再介绍更好的算法.

于是3.4节的第一积分对 $n=-1$ 也有了答案，由对数来填补，这是对积分表的突破. 由积分公式(3-18)和基本定理(3-7)，猜到对数的导数

$$(\ln x)'=\frac{1}{x}.$$

这容易由定义（2-2）来验证：在点 x（固定），

$$\left|\ln(x+h)-\ln(x)-\frac{1}{x}h\right|=\left|\int_x^{x+h}\left(\frac{1}{t}-\frac{1}{x}\right)\mathrm{d}t\right|$$
$$=\left|\int_x^{x+h}\frac{x-t}{xt}\mathrm{d}t\right|\leqslant\frac{1}{x^2}h^2.$$

这里用到积分区间的可加性，见3.6节.

由此可以导出对数函数的性质①，证明参看拉克斯等(1980) 及 G. Strang(1991). 把对数函数反过来，就可以定义指数函数：

$$\ln(e^x)=x,$$

求导后

$$(\ln(e^x))'=\frac{(e^x)'}{e^x}=1,$$

所以

$$(e^x)'=e^x.$$

这就是指数函数 e^x 令人吃惊之处：它的导数就是它自身，不变. 如果说别的函数求导不一定记得牢，那么 e^x 的求导不会记错. 由此，在积分表中又添了一道美丽的“风景”：

$$\int e^x dx = e^x.$$

更重要的是(Strang,1991)，指数函数解出了应用上最重要的一个微分方程：

$$f'(x)=f(x),\quad f(0)=1.$$

这也是法国数学课程标准（2001）中介绍的方程，详见第5章.

以上是 M. Livshits(2008)的路线，由我的学生陈竑焘具体计算的.

这是一个例子，开始由积分表的内在需要出发，引出自然对

---

① 单调递增；乘法的对数变成对数的加法：$\ln(ax)=\ln a+\ln x$；在0附近趋于 $-\infty$，在 ∞ 附近趋于 ∞.

数函数，它的反函数又引出了自然指数函数，结果解出了应用上最重要的微分方程（见 5.1.2 节：中国人口问题），这个解的实际算法见泰勒级数（4.4 节）. 所以，理论→应用→算法，或说纯数学、应用数学、计算数学成为一体.

对于其他还不在表中的函数，如 $\sqrt{1-x^2}$ 、$\ln x$，怎样求出它们的原函数，可以试用其他方法. 常用的有以下方法.

3.6 积分代换法

积分的加法和数乘由 2.5 节的微分法则容易得到．从略．积分对区间的可加性可由黎曼积分的定义得到（将区间中点取作分点）.

由 2.5 节的链式法则（iv）

$$(f(g(x)))'=f'(g(x))g'(x),$$

导出积分代换法

$$\int_a^b f'(g(x))g'(x)\,\mathrm{d}x=\int_a^b (f(g(x)))'\mathrm{d}x$$

$$=f(g(b))-f(g(a))=\int_{g(a)}^{g(b)} f'(x)\,\mathrm{d}x,$$

这使得在计算复杂函数的积分时，有时能变成 3.4 节积分表已有的函数，但使用时技巧性较强. 例如，倒过来用，令 $x=\sin t$，

$$\int_{-1}^{1}\sqrt{1-x^2}\,\mathrm{d}x=\int_{-\pi/2}^{\pi/2}\sqrt{1-\sin^2 t}\,\cos t\mathrm{d}t$$

$$=\int_{-\pi/2}^{\pi/2}\cos^2 t\mathrm{d}t=\frac{1}{2}\int_{-\pi/2}^{\pi/2}(1+\cos 2t)\,\mathrm{d}t.$$

后者能在3.4节积分表查到. 又如,

$$\int e^{kx}dx=\int e^t\frac{1}{k}dt=\frac{1}{k}\int e^t dt$$

(令 $kx=t, kdx=dt$),作为练习,由读者自己作上下限的变换;

$$\int\frac{xdx}{\sqrt{a^2-x^2}}=-\int dt$$

(令 $t=\sqrt{a^2-x^2}, dt=-xdx/\sqrt{a^2-x^2}$),作为练习,由读者自己作上下限的变换.

3.7 分部积分法

令 $f'(x)=u(x), g'(x)=v(x)$. 根据公式 (3-7),

$$\int u(x)dx=f(x),\quad \int v(x)dx=g(x),$$

有求和法则

$$\int(u+v)dx=\int(f+g)'dx=f(x)+g(x)$$

$$=\int udx+\int vdx,$$

根据公式(3-7)和乘积法则:

$$u(x)v(x)=\int(uv)'dx=\int(uv'+u'v)dx$$

$$=\int uv'dx+\int u'vdx$$

(假设其中一个积分存在),可以得到分部积分:

$$\int uv'dx=u(x)v(x)-\int u'vdx.$$

分部积分是一个小诀窍，关键是选择 u 和 v. 选择的目的是为了使 $\int u'v$ 比 $\int uv'$ 容易. 这可以从下面的例子看出.

例 $\int xe^x dx$ 不能在表中查到，试令 $u=x$ 和 $v'=e^x$，有 $u'=1$ 和 $v=e^x$，所以

$$\int xe^x dx = xe^x - \int 1 \cdot e^x dx = xe^x - e^x,$$

这里，我们把 xe^x 的积分转化为表中 e^x 的积分. 这里分部积分起的作用就是把幂函数的幂次降低一次.

又如，$\int x\cos x dx$，试令 $u=x$ 和 $v'=\cos x$，有 $u'=1$ 和 $v=\sin x$，所以

$$\int x\cos x dx = x\sin x - \int \sin x dx = x\sin x + \cos x,$$

这里，我们把 $x\cos x$ 的积分转化为表中 $\sin x$ 的积分. 这里分部积分起的作用仍是把幂函数的幂次降低一次.

再如，对数函数的积分 $\int \ln x dx$ 不能在表中查到，试令 $u=\ln x$ 和 $v'=1$，有 $u'=1/x$ 和 $v=x$，所以

$$\int \ln x dx = x\ln x - \int dx = x\ln x - x.$$

这里，我们把 $\ln x$ 的积分转化为数 1 的积分.

有些例子甚至需要多次使用分部积分公式：

$$\begin{aligned}\int e^x \sin x dx &= e^x \sin x - \int e^x \cos x dx \\ &= e^x \sin x - e^x \cos x - \int e^x \sin x dx\end{aligned}$$

或

$$\int e^x \sin x dx = \frac{1}{2}e^x(\sin x - \cos x).$$

可见，以上两种方法大大扩充了能够求出积分的范围.

注1　凡是能直接用基本公式的，不应把时间费在分部积分上. 例如，数学家华罗庚有一道考题：$\int f(x)f'(x)dx = ?$ 就应该直接用基本公式

$$\int f(x)f'(x)dx = \frac{1}{2}f^2(x),$$

这叫做基本功. 基本公式的潜力还表现在4.2节，那里连用几次基本公式(不用分部积分)便套出泰勒定理.

注2　积分法和微分法的处境完全不同. 对任意初等函数，它的导数总能计算出来(也是初等函数). 但存在着这样的初等函数，如 $\sin x/x$(在工程中)、e^{-x^2}(在概率论中)、$1/\ln x$、$\sin x^2$、$\sqrt{1+\cos^2 x}$(在弧长计算中)，它们的积分不再是初等函数. 事实上，具体的积分很少能通过(3-5)来计算. 要得到这些函数的积分，只能求助于数值方法，见第4章附录.

所以，(3-7)作为积分的计算不是完全成功的(卓里奇，2006)，它的作用首先在于把积分跟微分联系起来.

3.8　面积测量

把1.5节由白话翻译成函数.

积分值本来是当区间不断细分时微分求和的最终结果，或者想成斜率图的面积(图3-8). 这是无数矩形的面积的和. 这个算术本来难以完成，但是，由于有了等式(3-7)，一举取得最终结果，实在奥妙.

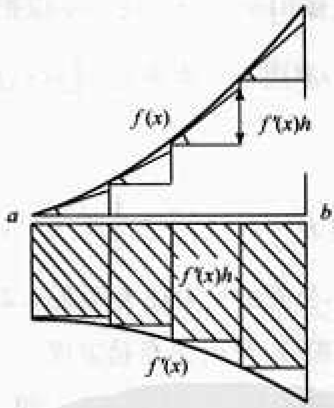


图3-8 高度图与斜率图

3.9 弧长测量

把(1-18)由白话翻译成函数：

$$s=\int_a^b\sqrt{1+u'(x)^2}\,dx. \tag{3-19}$$

这正是中学勾股公式的推广：由直到曲. 它的组成单位(微元)都是中学勾股公式(1-17). 所幸的是，曲和直之间的差别能消失(见第5章附录)，微元法能生效. 正如勾股公式能够用来计算斜边的长，但它首先是指出三边之间的关系. 同理，弧长公式(3-19)也是首先指出不同量之间的关系，当然也能用它来计算圆

的周长：求单位圆 $u(x)=\sqrt{1-x^2}$ 的半周长

$$\int_{-1}^{1}\sqrt{1+\left(\frac{-x}{\sqrt{1-x^2}}\right)^2}\,dx=\int_{-1}^{1}\frac{1-x^2+x^2}{\sqrt{1-x^2}}\,dx \qquad (3\text{-}20)$$

$$=\int_{-1}^{1}\sqrt{1-x^2}\,dx-\frac{1}{2}\int_{-1}^{1}\frac{x\,d(1-x)^2}{\sqrt{1-x^2}}$$

$$=\int_{-1}^{1}\sqrt{1-x^2}\,dx-\int_{-1}^{1}x\,d\sqrt{1-x^2}$$

$$=2\int_{-1}^{1}\sqrt{1-x^2}\,dx.$$

这里式（3-20）的被积函数是无界的，还需要另行处理，这里从略. 可见，用(3-19)来计算弧长可能很复杂. 它首先是说明弧长和斜率的关系.

阿诺德(2006)给出一个有说服力的例子，他指出：一个标准正弦曲线的长度：

$$\int_0^{2\pi}\sqrt{1+\cos^2 x}\,dx\approx 7.64.$$

只比横轴的长度(2π)大约增加20%(图3-9).

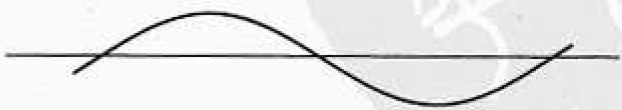


图3-9　正弦曲线之长

例如，5米的横轴大约相当于6米的标准正弦曲线. 直观只能说曲线比横轴长，只有弧长公式才能定量. 这里真正用到了斜率或导数.

注3　一般教材是用小段上的割线长来代替这段上的曲线长，

注 3 一般教材是用小段上的割线长来代替这段上的曲线长，然后又把割线转化为切线(因为我们只知道曲线上的斜率). 我们认为这反而引起麻烦(因为必须用中值定理)，不如直接用小段起点的切线长来代替小段曲线长，它并不改变最后的结果.

第 4 章
泰勒定理(基本定理的连用)

没有数学符号，泰勒不能展开.

基本公式(3-7)能够改写许多公式.

第 3 章由微分(2-2)(作为原因)导出基本定理(3-7)(作为结果). 现在颠倒过来，由结果(3-7)改写原因（微分).

4.1　重写微分定义

本来是由微分(2-2)或(2-8)导出积分(3-7). 现在颠倒过来，由积分(3-7)

$$f(x+h)-f(x)=\int_x^{x+h} f'(s)\,ds,$$

改写微分（2-8)：对于任一中间点 $\xi \in [x, x+h]$，

$$f(x+h)-f(x)-hf'(\xi)=\int_x^{x+h} [f'(s)-f'(\xi)]\,ds.$$

由此认出了相对误差的上界

$$\epsilon(h)=\underset{s,t\in[x,x+h],x\in[a,b]}{\text{upper}}|f'(s)-f'(t)|,$$

这里我们利用了一个事实，若$|g|\leqslant C$,则

$$\left|\int g(s)\mathrm{d}s\right|\leqslant C\left(\int \mathrm{d}s\right).$$

理由：由定义(2-2)出发，令$g=f'$,

$$\begin{aligned}\int_x^{x+h}g(s)\mathrm{d}s&=f(x+h)-f(x)\\&=f'(x)h+\epsilon(x,h)h\\&=g(x)h+\epsilon(x,h)h,\end{aligned}$$

相加得

$$\int_a^b g(s)\mathrm{d}s=\sum\int_x^{x+h}g(s)\mathrm{d}s=\sum[g(x)+\epsilon(x,h)]h.$$

再估计

$$\begin{aligned}\left|\int_a^b g(s)\mathrm{d}s\right|&\leqslant\sum[|g(x)|+|\epsilon(x,h)|]h\\&\leqslant[C+\epsilon(h)](b-a),\end{aligned}$$

其中用到假设(2-4)：$|\epsilon(x,h)|\leqslant\epsilon(h)\ll 1$. 于是只能$\left|\int_a^b g(s)\mathrm{d}s\right|\leqslant C(b-a)$.

4.2 重写泰勒余项

泰勒定理是微分(2-2)的延伸，它的真面目不过是基本定理(3-7)，连用几次. 下面采用缩写$a:\approx b$表示$|a|\leqslant b$.

首先，基本定理可以认为是函数 f 在点 x 的附近点 $x+h$ 被换成同一常数，$f(\xi)$ ($\xi\in[x,x+h]$)，所造成的误差：

$$f(x+h)-f(\xi)=\int_{\xi}^{x+h}f'(s)\,\mathrm{d}s\leqslant h\mathrm{Av}(f')\ :\approx\ h\underset{[x,x+h]}{\mathrm{upper}}|f'|$$

其中，$\mathrm{Av}(f')$ 表示 f' 的无限个数值的平均(或者避开它，跳到最后的结果).

微分(2-8)则认为是函数 f 在点 x 的附近被换成关于 h 的线性函数 $f(x)+hf'(\xi)$ 所造成的误差：

$$\begin{aligned}&f(x+h)-f(x)-hf'(\xi)\\&=\int_{x}^{x+h}[f'(s_2)-f'(\xi)]\,\mathrm{d}s_2=\int_{x}^{x+h}\int_{\xi}^{s_2}f''(s_1)\,\mathrm{d}s_1\mathrm{d}s_2\\&:\approx\underset{[x,x+h]}{\mathrm{upper}}|f''|\int_{x}^{x+h}\int_{x}^{x+h}\mathrm{d}s_1\mathrm{d}s_2=h^2\underset{[x,x+h]}{\mathrm{upper}}|f''|,\end{aligned}$$

这里的诀窍就是每逢相减变成一个积分(副产品：相对误差 $\approx Ch$，与点 x 无关，并随着 h 一起变小). 同理，f 在起点 x 附近能够继续换成关于 h 的高次函数，所造成的误差与 h^2，h^3，…相比也很小：

$$\begin{aligned}&f(x+h)-f(x)-hf'(x)-\frac{h^2}{2}f''(\xi)\\&=\int_{x}^{x+h}\int_{x}^{s_3}[f''(s_2)-f''(\xi)]\,\mathrm{d}s_2\mathrm{d}s_3=\int_{x}^{x+h}\int_{x}^{s_3}\int_{\xi}^{s_2}f'''(s_1)\,\mathrm{d}s_1\mathrm{d}s_2\mathrm{d}s_3\\&:\approx\underset{[x,x+h]}{\mathrm{upper}}|f'''|\int_{x}^{x+h}\int_{x}^{x+h}\int_{x}^{x+h}\mathrm{d}s_1\mathrm{d}s_2\mathrm{d}s_3=h^3\underset{[x,x+h]}{\mathrm{upper}}|f'''|,\qquad(4\text{-}1)\end{aligned}$$

$$\begin{aligned}&f(x+h)-f(x)-hf'(x)-\frac{h^2}{2}f''(x)-\frac{h^3}{3!}f'''(\xi)\\&:\approx h^4\underset{[x,x+h]}{\mathrm{upper}}|f^{(4)}|,\end{aligned}$$

$$f(x+h)-f(x)-hf'(x)-\frac{h^2}{2}f''(x)-\cdots-\frac{h^n}{n!}f^{(n)}(\xi)$$

$$:\approx h^{n+1}\underset{[x,x+h]}{\text{upper}}|f^{(n+1)}|.$$

这是具有累次积分余项的泰勒定理.

下面改 x 为 x_0、$x+h$ 为 x，并讨论余项的求导. 取出余项，如(4-1)中的

$$R(x)=\int_{x_0}^{x}\int_{x_0}^{s_1}\int_{x_0}^{s_2}f'''(s_3)\,\mathrm{d}s_3\mathrm{d}s_2\mathrm{d}s_1,\quad |R|\leqslant\frac{h^3}{6}\text{upper}|f'''|,$$

它的导数为

$$R'(x)=\int_{x_0}^{x}\int_{x_0}^{s_2}f'''(s_3)\,\mathrm{d}s_3\mathrm{d}s_2,\quad |R'|\leqslant\frac{h^2}{2}\text{upper}|f'''|,$$

$$R''(x)=\int_{x_0}^{x}f'''(s_3)\,\mathrm{d}s_3,\quad |R''|\leqslant h\ \text{upper}\ |f'''|\ ,$$

$$R'''(x)=f'''(x),$$

其中余项的各阶导数(又称全范数)只用高阶导数(又称半范数)来控制，这就是计算数学误差分析的理论根据：Bramble-Hilbert-Xu 引理.

刘嘉荃指出余项也可回到单重积分形式(需要小技巧，从略).

注 经典的泰勒余项是由陆续作分部积分来证明的，这需要多花时间和脑筋(技巧)，难记，不妨称为间接方法. 我们只记住基本定理，急中生智，就直接用基本定理，连用几次就套出来了，多轻松！这叫做因祸得福，也称为直接方法.

泰勒展开式的应用能取材于许多教材，例如，华罗庚(1979)、G. Strang(1991)、Y. Zeldovich-I. Yaglom(1987)等的著作.

4.3 数值积分

学生以为数值积分很复杂，其实就是泰勒展开：视 $f' = g$，则泰勒展开就是小段上积分的一种求积公式．例如，

$$\left|\int_x^{x+h} g(s)\,\mathrm{d}s - hg(x)\right| \leqslant \frac{h^2}{2}\underset{[x,x+h]}{\text{upper}}|g'|,$$

$$\left|\int_x^{x+h} g(s)\,\mathrm{d}s - hg(x) - \frac{h^2}{2}g'(x)\right| \leqslant \frac{h^3}{6}\underset{[x,x+h]}{\text{upper}}|g''|.$$

梯形、中点、辛普森等求积公式不过是上式的变形，见附录．

4.4 泰勒级数

在上面的泰勒展开中，改 x 为 0，$x+h$ 为 x，便得

$$f(x) - f(0) = f'(0)x + \cdots + \frac{1}{n!}f^{(n)}(0)x^n + R,$$

$$|R| \leqslant \frac{x^{n+1}}{(n+1)!}\,\text{upper}\,|f^{(n+1)}|.$$

如果当 n 充分大后，$f^{(n+1)}$ 都是有界的，则整个无限和是一个泰勒级数．所以很多函数可以展开为级数，特别值得注意的是下列函数的展开：

$$\mathrm{e}^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots,$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots,$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots,$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots,$$

$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \cdots$ （其中 n 可以是有理数,甚至无理数),

其中第一式特别好记，且收敛很快. 以 $x=1$ 代入，得

$$e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!} + \cdots,$$

这个 e 的近似值是

$$e = 2.718281828459045235 36\cdots.$$

4.5 欧拉公式

把 e^{ix} 定义为一个级数

$$e^{ix} = 1 + \frac{1}{1!}(ix) + \frac{1}{2!}(ix)^2 + \frac{1}{3!}(ix)^3 + \frac{1}{4!}(ix)^4 + \frac{1}{5!}(ix)^5 + \cdots,$$

利用

$$i^2 = -1$$

把级数分为奇数项和偶数项，认出为余弦和正弦：

$$\begin{aligned} e^{ix} &= \left(1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \cdots\right) + i\left(\frac{1}{1!}x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots\right) \\ &= \cos x + i\sin x. \end{aligned}$$

它与 e^x(指数增长)截然不同，含有振动的分量. 如果前者用于生命科学、经济学(所谓一阶微分方程),后者则用于物理、工程

(所谓二阶微分方程). 见第 5 章.

进一步地,

$$(e^{ix})' = -\sin x + i\cos x = i(\cos x + i\sin x) = ie^{ix}.$$

同理, 对于一个复数 λ, 有

$$(e^{\lambda x})' = \lambda e^{\lambda x}, \quad (e^{\lambda x})'' = \lambda^2 e^{\lambda x}$$

等. 也就是函数 $e^{\lambda x}$ 的导数, 保持了同样的类型(只有系数不同). 这经常用来解二阶微分方程. 也见第 5 章.

附录　积分数值方法

级数展开式可以省去一项或几项, 便得著名的求积公式: 在每一子区间上, 有

梯形公式:

$$\int_x^{x+h} g(s_2)\,ds_2 - \frac{h}{2}[g(x) + g(x+h)] :\approx \frac{5h^3}{12} \underset{[x,x+h]}{\text{upper}} |g''|.$$

中点公式:

$$\int_x^{x+h} g(s_3)\,ds_3 - hg\left(x + \frac{h}{2}\right) :\approx \frac{7h^3}{24} \underset{[x,x+h]}{\text{upper}} |g''|,$$

$$\int_x^{x+h} g(s_4)\,ds_4 - hg(x) - \frac{h^2}{2} g'\left(x + \frac{h}{3}\right) :\approx \frac{5h^4}{72} \underset{[x,x+h]}{\text{upper}} |g'''|,$$

$$\int_x^{x+h} g(s_5)\,ds_5 - \frac{h}{2}[g(x) + g(x+h)] + \frac{h^3}{12} g''\left(x + \frac{h}{2}\right)$$

$$:\approx \frac{19h^5}{480} \underset{[x,x+h]}{\text{upper}} |g^{(4)}|,$$

辛普森公式 $= \dfrac{2 \times \text{中点公式} + \text{梯形公式}}{3} :\approx \dfrac{49h^5}{2880} \underset{[x,x+h]}{\text{upper}} |g^{(4)}|$, 这

里，缩写 $a:\approx b$ 是代表 $|a|\leqslant b$. 其证明只是演算，太长了，受不了（除非有简化证明），不必读，仅作为备忘录：首先，

$$
\begin{aligned}
&[g(x)+g(x+h)]-2g(x)-hg'(x)\\
&=\int_x^{x+h} g'(s_3)\,\mathrm{d}s_3-hg'(x):\approx \frac{h^2}{2}\underset{[x,x+h]}{\text{upper}}|g''|,\\
&g\left(x+\frac{h}{2}\right)-g(x)-\frac{h}{2}g'(x)\\
&=\int_x^{x+\frac{h}{2}} g'(s_2)\,\mathrm{d}s_2-\frac{h}{2}g'(x):\approx \frac{h^2}{8}\underset{[x,x+h]}{\text{upper}}|g''|
\end{aligned}
$$

（其中第一式中的 h 改为 $h/2$）. 其次，把 g 变为 g'，同时把 $h/2$ 变为 $h/3$，

$$
g'\left(x+\frac{h}{3}\right)-g'(x)-\frac{h}{3}g''(x):\approx \frac{h^2}{18}\underset{[x,x+h]}{\text{upper}}|g'''|.
$$

最后，

$$
\begin{aligned}
&[g(x)+g(x+h)]-\frac{h^2}{6}g''\left(x+\frac{h}{2}\right)-2g(x)-hg'(x)\\
&-\frac{h^2}{3}g''(x)-\frac{h^3}{12}g'''(x)\\
&=\left[g(x)+g(x+h)-2g(x)-hg'(x)-\frac{h^2}{2}g''(x)-\frac{h^3}{6}g'''(x)\right]\\
&-\frac{h^2}{6}\left[g''\left(x+\frac{h}{2}\right)-g''(x)-\frac{h}{2}g''(x)\right]\\
&=\left[\int_x^{x+h}\int_x^{s_3} g''(s_4)\,\mathrm{d}s_4\mathrm{d}s_3-\frac{h^2}{2}g''(x)-\frac{h^3}{6}g'''(x)\right]\\
&-\frac{h^2}{6}\left[\int_x^{x+h} g'''(s_5)\,\mathrm{d}s_5-\frac{h}{2}g''(x)\right]
\end{aligned}
$$

$$= \left[\int_x^{x+h}\int_x^{s_3}\int_x^{s_4} g'''(s_5)\,ds_5 ds_4 ds_3 - \frac{h^3}{6}g'''(x)\right]$$

$$- \frac{h^2}{6}\int_x^{x+h}\int_x^{s_5} g^{(4)}(s_6)\,ds_6 ds_5$$

$$= \int_x^{x+h}\int_x^{s_3}\int_x^{s_4}\int_x^{s_5} g^{(4)}(s_6)\,ds_6 ds_5 ds_4 ds_3$$

$$- \frac{h^2}{6}\int_x^{x+h}\int_x^{s_5} g^{(4)}(s_6)\,ds_6 ds_5$$

$$:\approx \frac{h^4}{24}\underset{[x,x+h]}{\text{upper}}|g^{(4)}| + \frac{h^4}{48}\underset{[x,x+h]}{\text{upper}}|g^{(4)}|$$

$$= \frac{h^4}{16}\underset{[x,x+h]}{\text{upper}}|g^{(4)}|,$$

$$\frac{2g(x+h/2) + 1/2[g(x)+g(x+h)]}{3} - g(x) - \frac{h}{2}g'(x)$$

$$- \frac{h^2}{3!}g''(x) - \frac{h^3}{4!}g'''(x)$$

$$= \frac{2}{3}\left[g\left(x+\frac{h}{2}\right) - g(x) - \frac{h}{2}g'(x) - \frac{h^2}{8}g''(x) - \frac{h^3}{48}g'''(x)\right]$$

$$+ \frac{1}{6}\left[g(x) + g(x+h) - 2g(x) - hg'(x) - \frac{h^2}{2}g''(x) - \frac{h^3}{6}g'''(x)\right]$$

$$= \frac{2}{3}\left[\int_x^{x+h/2}\int_x^{s_2} g''(s_3)\,ds_3 ds_2 - \frac{h^2}{8}g''(x) - \frac{h^3}{48}g'''(x)\right]$$

$$+ \frac{1}{6}\left[\int_x^{x+h}\int_x^{s_3} g''(s_4)\,ds_4 ds_3 - \frac{h^2}{2}g''(x) - \frac{h^3}{6}g'''(x)\right]$$

$$= \frac{2}{3}\left[\int_x^{x+h/2}\int_x^{s_2}\int_x^{s_3} g'''(s_4)\,ds_4 ds_3 ds_2 - \frac{h^3}{48}g'''(x)\right]$$

$$+ \frac{1}{6}\left[\int_x^{x+h}\int_x^{s_2}\int_x^{s_3} g'''(s_4)\,ds_4 ds_3 ds_2 - \frac{h^3}{6}g'''(x)\right]$$

$$= \frac{2}{3}\int_x^{x+h/2}\int_x^{s_2}\int_x^{s_3}\int_x^{s_4} g^{(4)}(s_5)\,\mathrm{d}s_5\mathrm{d}s_4\mathrm{d}s_3\mathrm{d}s_2$$

$$+ \frac{1}{6}\int_x^{x+h}\int_x^{s_2}\int_x^{s_3}\int_x^{s_4} g^{(4)}(s_5)\,\mathrm{d}s_5\mathrm{d}s_4\mathrm{d}s_3\mathrm{d}s_2$$

$$:\approx \frac{h^4}{576}\underset{[x,x+h]}{\text{upper}}|g^{(4)}| + \frac{h^4}{144}\underset{[x,x+h]}{\text{upper}}|g^{(4)}| = \frac{5h^4}{576}\underset{[x,x+h]}{\text{upper}}|g^{(4)}|.$$

因此,矩形、中点、辛普森等公式成立,即

$$\int_x^{x+h} g(s_2)\,\mathrm{d}s_2 - \frac{h}{2}[g(x)+g(x+h)]$$

$$= \int_x^{x+h} g(s_2)\,\mathrm{d}s_2 - hg(x) - \frac{h^2}{2}g'(x)$$

$$- \frac{h}{2}[g(x)+g(x+h)-2g(x)-hg'(x)]$$

$$:\approx \frac{h^3}{3!}\underset{[x,x+h]}{\text{upper}}|g''| + \frac{h^3}{4}\underset{[x,x+h]}{\text{upper}}|g''| = \frac{5h^3}{12}\underset{[x,x+h]}{\text{upper}}|g''|,$$

$$\int_x^{x+h} g(s_3)\,\mathrm{d}s_3 - hg\left(x+\frac{h}{2}\right)$$

$$= \int_x^{x+h} g(s_3)\,\mathrm{d}s_3 - hg(x) - \frac{h^2}{2}g'(x)$$

$$- h\left[g\left(x+\frac{h}{2}\right) - g(x) - \frac{h}{2}g'(x)\right]$$

$$:\approx \frac{h^3}{3!}\underset{[x,x+h]}{\text{upper}}|g''| + \frac{h^3}{8}\underset{[x,x+h]}{\text{upper}}|g''| = \frac{7h^3}{24}\underset{[x,x+h]}{\text{upper}}|g''|,$$

$$\int_x^{x+h} g(s_4)\,\mathrm{d}s_4 - hg(x) - \frac{h^2}{2}g'\left(x+\frac{h}{3}\right)$$

$$= \int_x^{x+h} g(s_4)\,\mathrm{d}s_4 - hg(x) - \frac{h^2}{2}g'(x) - \frac{h^3}{6}g''(x)$$

$$- \frac{h^2}{2}\left[g'\left(x+\frac{h}{3}\right) - g'(x) - \frac{h}{3}g''(x)\right]$$

$$:\approx \frac{h^4}{4!}\underset{[x,x+h]}{\text{upper}}|g'''| + \frac{h^4}{18}\underset{[x,x+h]}{\text{upper}}|g'''| = \frac{7h^4}{72}\underset{[x,x+h]}{\text{upper}}|g'''|,$$

$$\int_x^{x+h} g(s_5)\,ds_5 - \frac{h}{2}\left[g(x)+g(x+h)\right] + \frac{h^3}{12}g''\left(x+\frac{h}{2}\right)$$

$$= \int_x^{x+h} g(s_5)\,ds_5 - hg(x) - \frac{h^2}{2}g'(x) - \frac{h^3}{3!}g''(x) - \frac{h^4}{4!}g'''(x)$$

$$-\frac{h}{2}\Big[g(x)+g(x+h) - \frac{h^2}{6}g''\left(x+\frac{h}{2}\right) - 2g(x)$$

$$-hg'(x) - \frac{h^2}{3}g''(x) - \frac{h^3}{12}g'''(x)\Big]$$

$$:\approx \frac{h^5}{5!}\underset{[x,x+h]}{\text{upper}}|g^{(4)}(x)| + \frac{h^5}{32}\underset{[x,x+h]}{\text{upper}}|g^{(4)}| = \frac{19h^5}{480}\underset{[x,x+h]}{\text{upper}}|g^{(4)}|,$$

$$\int_x^{x+h} g(s_2)\,ds_2 - \frac{2g(x+h/2) + 1/2[g(x)+g(x+h)]}{3}$$

$$= \int_x^{x+h} g(s_2)\,ds_2 - hg(x) - \frac{h^2}{2}g'(x) - \frac{h^3}{3!}g''(x) - \frac{h^4}{4!}g'''(x)$$

$$+h\Big[\frac{2g(x+h/2) + 1/2[g(x)+g(x+h)]}{3}$$

$$-g(x) - \frac{h}{2}g'(x) - \frac{h^2}{3!}g''(x) - \frac{h^3}{4!}g'''(x)\Big]$$

$$:\approx \frac{h^5}{5!}\underset{[x,x+h]}{\text{upper}}|g^{(4)}| + \frac{5h^5}{576}\underset{[x,x+h]}{\text{upper}}|g^{(4)}| = \frac{49h^5}{2880}\underset{[x,x+h]}{\text{upper}}|g^{(4)}|.$$ ①

① 以上是我的学生谢和虎、贾尚辉帮助算的(应该还可以简化和改进).

第5章 微分方程(才用到实数)

前4章无需实数，也能拿下；本章如无实数，便拿不下.

有两个最简单的微分方程，第一个能由前面有过的基本定理(3-7)直接求解，第二个不能由基本定理求解，需要利用3.5节造出的指数函数求解. 然后，其他方程都要变成这两个方程(所谓母方程)才能解出.

法国数学课程标准(2001)高三的微积分也介绍了这些方程(但无证明)，不能说太难.

5.1 一阶微分方程

5.1.1 第一方程

前面见过的基本定理(3-7)，给出了高和斜率之间的关系式：

已知山坡f的斜率

$$f'(x)=g(x). \tag{5-1}$$

求山高f自身（$x\geqslant a$）. 这就是最简单的微分方程，有着最明显的几何说明，它有显式解

$$f(x)=f(a)+\int_a^x g(s)\,\mathrm{d}s. \tag{5-2}$$

它是一个函数. 但是中学只学代数方程，它的解只是一个数.

微分方程这个工具是牛顿创造的，用于经典力学. 此后就被许多科学家所继承，用到不同学科之中. 例如，马尔萨斯把它用于人口学，麦克斯韦把它用于电磁学，薛定谔把它用于量子力学. 这些微分方程对我们的生活有切身的影响，如人口预测. 手机以及纳米都和这些方程有联系.

5.1.2 第二方程

人口问题造出另一个方程，人口增长率f'依赖于当时的人口总量f:

$$f'(x)=rf(x). \tag{5-3}$$

法国数学课程标准(2001)高三的微积分就介绍(5-3)给中学生. G. Strang(1991)也曾指出，(5-3)是应用数学(生命科学、经济学等)中最重要的微分方程，数学家常用到它. 它与第一方程(5-1)不同，不能由基本公式(5-2)直接求解. 但是又必须造出一个解，否则许多方程也解不出来. 所以第二方程是一个关卡!

前面3.5节已造出(5-3)的一个解，表示为指数函数

$$f(x)=C\cdot e^{rx} \tag{5-4}$$

（如果 r 是变系数，则为 $C\cdot e^{\int r(x)\mathrm{d}x}$）．这是微积分（基本定理）的突破．它有什么意义呢？

这里讲一件发生在我国的实例．2000 年全国人口普查，挨家挨户实查一年多（约 5 亿人年），查出是 12.66 亿人，若用微分方程，一个大学生只花 5 分钟（5 人分），算出是 13.45 亿人．两种结果相差不多，但效率有天壤之别，前者又费时又慢（5 亿人年），后者又省又快（5 人分）．这是什么道理呢？后一方法的基本原理是：设 $f(x)$ 在时间 x 的人口数，于是根据马尔萨斯定律（5-3）来算：已知其中增长率 $r=0.0148$ 以及过去的数据 $f(1990)=11.6$ 亿人（朱学贤等，2000），根据（5-4）或（5-10），得出 $f(2000)=13.45$ 亿人．可见，数学能够作为预测或验证现实的一种又省又快的方法．说微分方程关系到国计民生的大事，并非天方夜谭．这就是为什么牛顿－莱布尼茨说"必须解微分方程"的道理．

注 1　刘嘉荃说，实查和预测相差 8000 万人（占 6.4%）；但他又指出，这可解释为人口流动和少报所造成．

与人口预测相似的，还有气象预报．在此引用吴文俊（1995）的描述："所谓气象预报，无非是根据过去一段时期，对各地压力、温度、降雨量等的实测数据，以及表达气象变化规律的这些函数间的微分方程，用数学方法推算出今后一段时期内的这些函数数值，以预报气象特征而已．"

举一反三，应有核爆炸预报．根据可测的数据，以及表达核

爆炸规律的微分方程，应该能由少数人在计算机上用数值计算方法，代替直接的大气或地下核试验，推算出核爆炸特征. 这就是现在流行的说法：科学工程计算已成为除理论和实验之外的第三手段.

以下，把第一方程(5-1)和第二方程(5-3)作为求解许多方程的母方程，别的方程都要变成这两个方程才能解出.

5.1.3 线性微分方程(分解为母方程)

考虑比第二方程(5-3)更一般的线性微分方程

$$f' + r(x)f = g.$$

法国数学课程标准(2001)高三的微积分也把它介绍给中学生.

注意，到此我们只有第一方程(5-1)和第二方程(5-3)可解，所以别的方程只有变成这两个方程时才可解. 为此，在上述线性方程两边乘以待定函数$\beta(x)$，

$$\beta f' + r(x)\beta f = \beta g,$$

使左边能配成一个导数：

$$\begin{cases} (\beta f)' = \beta f' + \beta' f, \\ \beta' = r(x)\beta, \end{cases}$$

其中，前者来自第一方程，后者来自第二方程，认出

$$\begin{cases} \beta f = \int (\beta f)' \mathrm{d}x = \int \beta g \, \mathrm{d}x, \\ \beta(x) = \mathrm{e}^{\int r(x)\mathrm{d}x}. \end{cases}$$

(第一式可以相差一个常数).

注 2 解方程有一点技巧. 中学解二次方程要配成平方, 这里解微分方程则配成导数.

5.1.4 分离变量方程

考虑方程右端能分离变量:

$$f'=\frac{g(x)}{l(f)},$$

两边乘以 $l(f)$:

$$l(f)f'=g(x),$$

使左边能配成一个导数:

$$\begin{cases}(G[f(x)])'=G'(f)f',\\ G'(f)=l(f)\end{cases}$$

(其中用到复合函数求导), 两者来自第一方程, 认出

$$\begin{cases}G[f(x)]=\int g(x)\,\mathrm{d}x,\\ G(f)=\int l(f)\,\mathrm{d}f\end{cases}$$

(第一式可以相差一个常数). 以上两个方程的解法, 见布朗(1978)《微分方程及其应用》.

5.1.5 更一般的方程

5.1.4 小节已经解出一阶方程 $f'=g(x)/l(f)$. 对于更一般的一阶方程

$$f'=g(x,f), \tag{5-5}$$

右端变量不能分离，我们便不能配成导数，不能有显式解．数学一方面转向(5-5)的存在和惟一性定理也见布朗(1978)；另一方面也转向(5-5)的近似解法，欧拉折线算法(法国数学课程标准(2001)高三的微积分也把它介绍给中学生)．为了更好理解这个算法，先对最简单的第一方程和第二方程做试探，看看它们是什么样的？跟显式解如何相像？

5.1.6　第一方程的欧拉算法

回望基本公式(3-7)的前身，见图5-1.

图5-1中不连续的折线，通过平移，能替换为连续折线，见图5-2.

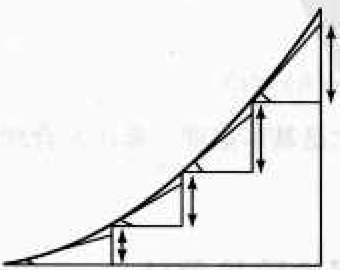


图5-1　求高图

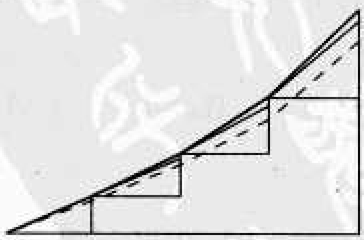


图5-2　连续折线

这两种折线的求高公式是一样的，都是基于微分或小段切线的高，是基本公式的有限步计算，能写为欧拉算法，见图5-3.

按(5-1)：$f'=g$，有

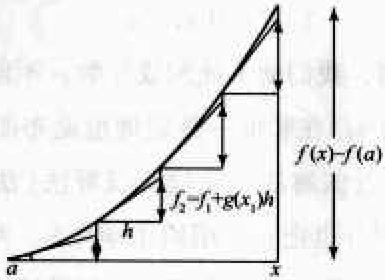


图 5-3 欧拉折线求高法

$$f_1 = f_0 + g(x_0)h,\quad f_{n+1} = f_n + g(x_n)h, \tag{5-6}$$

其中 $a = x_0 < x_1 < \cdots < x_n < x_{n+1} = x$，$h = x_{i+1} - x_i$，$f_0 = f(a)$，或写为

$$f_{n+1} = f(a) + \sum_{0 \leqslant i < n+1} g(x_i)h,$$

与第一方程的解基本定理（3-7）相像，是它的有限类比．有误差

$$E_{n+1} = f(x) - f_{n+1} = f(x) - f(a) - \sum_{0 \leqslant i < n+1} g(x_i)h$$

（此即 1.2 节的全段误差）．由（3-3），

$$|E_{n+1}| \leqslant (x-a)\epsilon(h) \leqslant Ch,$$

所以，第一方程(5-1)的欧拉算法就是基本定理，算法是合理的．

5.1.7 第二方程的欧拉算法

回到人口方程，又称第二方程

$$f'(x) = f(x)\ (x > 0),\quad f(0) = 1 \tag{5-7}$$

由于 f 是未知的，欧拉算法（5-6）中的已知数 $g(x_n)$ 不能用了，要改为 f_n：

$$f_{n+1}=(1+h)f_n=(1+h)^{n+1}f_0$$
$$=(1+h)^{\frac{x}{h}} \tag{5-8}$$

其中假设$f_0=1$，$(n+1)h=x$. 利用第二方程(5-7)，$f=f'$，能够推导误差

$$f_1=(1+h)f_0,$$
$$|f(x_1)-f_1|=|f(x_1)-f_0-f_0h|=\left|\int_0^h\int_0^s f(t)\,\mathrm{d}t\mathrm{d}s\right|$$
$$\leqslant\frac{1}{2}\max_{[0,x]}|f|\ h^2:=Ch^2,$$
$$f_2=(1+h)f_1,$$
$$|f(x_2)-f_2|$$
$$=|f(x_2)-f(x_1)-f(x_1)h+[f(x_1)-f_1](1+h)|$$
$$\leqslant Ch^2+Ch^2(1+h),$$

……

$$f_{n+1}=(1+h)f_n,$$
$$|f(x)-f_{n+1}|$$
$$=|f(x)-f(x_n)-f(x_n)h+[f(x_n)-f_n](1+h)|$$
$$\leqslant Ch^2+Ch^2(1+h)+\cdots+Ch^2(1+h)^n \quad (\text{归纳法})$$
$$=Ch[(1+h)^{n+1}-1] \quad (\text{等比数列求和}) \tag{5-9}$$
$$=Ch[(1+h)^{\frac{x}{h}}-1]$$
$$\leqslant Ch(3^x-1), \tag{5-10}$$

最后的不等式用到$(1+h)^{1/h}<3$（对一切h）. 令$h=1/n$，

$$\left(1+\frac{1}{n}\right)^n\leqslant 1+1+\frac{1}{2!}+\frac{1}{3!}+\cdots+\frac{1}{n!}$$

$$<1+1+\frac{1}{2}+\frac{1}{4}+\cdots+\frac{1}{2^{n-1}}<3$$

（华罗庚,1963）. 由(5-10)，欧拉序列(5-8)，f_{n+1}，或$(1+h)^{x/h}\to f(x)$，当$h\to 0$时；或$\left(1+\frac{x}{n+1}\right)^{n+1}\to f(x)$，当$n\to\infty$时，其中$f(x)$即第二方程的解$e^x$，由欧拉序列造出

$$e^x=\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n. \tag{5-11}$$

所以，欧拉算法也造出指数函数（5-11）.

注 3 回到第二方程（5-3），有了指数函数. 按刘嘉荃的建议，两边乘以e^{-rx}，有

$$e^{-rx}f'(x)=re^{-rx}f(x),$$

使它配成导数的形式

$$\frac{d}{dx}(e^{-rx}f(x))=0$$

（其中用到微分法则），这样才能解出

$$e^{-rx}f(x)=c$$

或

$$f(x)=ce^{rx}.$$

经过对第二方程（5-3）试探之后，实已包含一般方程（5-5）的欧拉算法

$$f_{n+1}=f_n+g(x_n,f_n)h_n,\quad f_0=f(a).$$

所以，G. Strang(1991)说：近代数学是精确公式和近似计算的结合，两者不可偏废.

5.2　二阶微分方程(牛刀杀鸡)

生命科学、经济学引出一阶微分方程(变化依赖于现状)，见 5.1 节. 物理、工程(振动和波动现象)引出二阶微分方程(牛顿定律是加速度，不是速度)，如初值问题

$$\begin{cases} f''+fr=0, \\ f(a)=f_0,\quad f'(a)=v_0, \end{cases} \tag{5-12}$$

其中，实数 $r>0$. 由于我们只知道一阶方程，若引入新变量

$$\begin{cases} v=f', \\ v'=-fr, \end{cases}$$

并采用矩阵记号(华罗庚,1963 年以来)，则二阶方程(5-12)能变成一阶的矩阵(两行两列)微分方程

$$\begin{cases} (f,v)'=(f,v)\begin{pmatrix} 0 & -r \\ 1 & 0 \end{pmatrix}, \\ (f,v)(a)=(f_0,v_0), \end{cases} \tag{5-13}$$

与一行一列的微分方程(5-3)和指数解(5-4)相类比，这个矩阵微分方程(5-13)应该有一个两行两列的指数解

$$(f,v)=(f_0,v_0)\mathrm{e}^{\begin{pmatrix} 0 & -r \\ 1 & 0 \end{pmatrix}x},$$

其中右端矩阵定义为级数（华罗庚，1963 年以来）

$$\mathrm{e}^{\begin{pmatrix} 0 & -r \\ 1 & 0 \end{pmatrix}x}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}+\frac{x}{1!}\begin{pmatrix} 0 & -r \\ 1 & 0 \end{pmatrix}+\frac{x^2}{2!}\begin{pmatrix} 0 & -r \\ 1 & 0 \end{pmatrix}^2+\frac{x^3}{3!}\begin{pmatrix} 0 & -r \\ 1 & 0 \end{pmatrix}^3+\cdots, \tag{5-14}$$

但是如何认出振动的分量呢？回顾4.5节的欧拉公式，当时利用

$$i^2=-1,$$

把 e^{ix} 的级数分为奇数项和偶数项，从而认出余弦和正弦

$$\begin{aligned}e^{ix}&=1+\frac{1}{1!}(ix)+\frac{1}{2!}(ix)^2+\frac{1}{3!}(ix)^3+\frac{1}{4!}(ix)^4+\frac{1}{5!}(ix)^5+\cdots\\&=\left(1-\frac{1}{2!}x^2+\frac{1}{4!}x^4-\cdots\right)+i\left(\frac{1}{1!}x-\frac{1}{3!}x^3+\frac{1}{5!}x^5-\cdots\right)\\&=\cos x+i\sin x.\end{aligned}$$

照猫画虎，现在应该利用

$$\begin{pmatrix}0&-r\\1&0\end{pmatrix}^2=-\begin{pmatrix}r&0\\0&r\end{pmatrix},$$

也把（5-14）分为奇数项和偶数项

$$\begin{aligned}e^{\begin{pmatrix}0&-r\\1&0\end{pmatrix}x}&=\left(\begin{pmatrix}1&0\\0&1\end{pmatrix}-\frac{x^2}{2!}\begin{pmatrix}r&0\\0&r\end{pmatrix}+\cdots\right)\\&\quad+\begin{pmatrix}0&-r\\1&0\end{pmatrix}\left(\frac{x}{1!}\begin{pmatrix}1&0\\0&1\end{pmatrix}-\frac{x^3}{3!}\begin{pmatrix}r&0\\0&r\end{pmatrix}+\cdots\right),\end{aligned}\qquad(5\text{-}15)$$

右边第一项认出为余弦函数

$$\begin{pmatrix}\cos(xr^{\frac{1}{2}})&0\\0&\cos(xr^{\frac{1}{2}})\end{pmatrix}.$$

右边第二项认出为正弦函数

$$\begin{aligned}&\begin{pmatrix}0&-r\\1&0\end{pmatrix}\left(\frac{x}{1!}\begin{pmatrix}1&0\\0&1\end{pmatrix}-\frac{x^3}{3!}\begin{pmatrix}r&0\\0&r\end{pmatrix}+\cdots\right)\\&=\begin{pmatrix}0&-r\\1&0\end{pmatrix}\begin{pmatrix}r^{-\frac{1}{2}}&0\\0&r^{-\frac{1}{2}}\end{pmatrix}\left(\frac{x}{1!}\begin{pmatrix}r^{\frac{1}{2}}&0\\0&r^{\frac{1}{2}}\end{pmatrix}-\frac{x^3}{3!}\begin{pmatrix}r^{\frac{3}{2}}&0\\0&r^{\frac{3}{2}}\end{pmatrix}+\cdots\right)\end{aligned}$$

$$=\begin{pmatrix}0 & -r^{\frac{1}{2}}\\ r^{-\frac{1}{2}} & 0\end{pmatrix}\begin{pmatrix}\sin(xr^{\frac{1}{2}}) & 0\\ 0 & \sin(xr^{\frac{1}{2}})\end{pmatrix}$$

$$=\begin{pmatrix}0 & -r^{\frac{1}{2}}\sin(xr^{\frac{1}{2}})\\ r^{-\frac{1}{2}}\sin(xr^{\frac{1}{2}}) & 0\end{pmatrix}.$$

最后，(5-15) 认出了振动解

$$e^{\begin{pmatrix}0 & -r\\ 1 & 0\end{pmatrix}x}=\begin{pmatrix}\cos(xr^{\frac{1}{2}}) & -r^{\frac{1}{2}}\sin(xr^{\frac{1}{2}})\\ r^{-\frac{1}{2}}\sin(xr^{\frac{1}{2}}) & \cos(xr^{\frac{1}{2}})\end{pmatrix}.$$

最后得到

$$(f,v)=(f_0,v_0)\begin{pmatrix}\cos(xr^{\frac{1}{2}}) & -r^{\frac{1}{2}}\sin(xr^{\frac{1}{2}})\\ r^{-\frac{1}{2}}\sin(xr^{\frac{1}{2}}) & \cos(xr^{\frac{1}{2}})\end{pmatrix}$$

或

$$\begin{cases}f=f_0\cos(xr^{\frac{1}{2}})+v_0r^{-\frac{1}{2}}\sin(xr^{\frac{1}{2}}),\\ v=v_0\cos(xr^{\frac{1}{2}})-f_0r^{\frac{1}{2}}\sin(xr^{\frac{1}{2}}).\end{cases}$$

虽然用矩阵解法. 绕得太远了，但得到的这个显式解，能直接验证它的确满足方程(5-12). 更重要的，矩阵法适用于大方程组!

刘嘉荃说，没有必要绕这个圈子，直接猜一下：

$$(\sin xr^{\frac{1}{2}})''=-r\sin(xr^{\frac{1}{2}}),$$

$$(\cos xr^{\frac{1}{2}})''=-r\cos(xr^{\frac{1}{2}}).$$

对于一般的二阶微分方程

$$pf''+qf'+rf=0,$$

它的解应该形如 $e^{\lambda x}$，其中 λ 满足

$$p\lambda^2+q\lambda+r=0.$$

它有两个根：二实、二复或实的重根，相应于增长、衰减和

振动.

5.3 后 记

从确定性走向随机，前者是后者的最初逼近.

G. Papanicolaou(1973)指出，在物理、工程、经济、生物和其他科学问题中提炼出来的数学模型几乎都是随机而不是确定的. 人们更关注的是事件发生的概率而不是每个偶然事件. 绝大多数的数学方法都是基于确定性的模型，但是变成确定性的模型是对随机问题的最初逼近，而且也是合理的. 以上探讨的是具有确定系数的二阶常微分方程，但是能够将其推广到随机系数的常微分方程.

另外，二阶方程还有边值问题，其显式解很麻烦，多数人受不了，不如暂时放弃，有余力者可读布朗(1978)的《微分方程及其应用》.

附录 微分方程的存在定理(初学略去)

第一方程的存在定理

回到最简单的微分方程：

$$f'(x)=g(x),\quad x\in[a,b];\ f(a)=\alpha \tag{1}$$

问题(文丽,1981)：对一个给定的函数 g，上述方程是否存

在解 f？当解存在，是否惟一？

有标准的答案：

当 g 一致连续时，

$$\underset{[a,b]}{\text{upper}}|g(x+h)-g(x)|\leqslant\epsilon_1(h)\ll 1 \tag{2}$$

（当 h 变小）方程(1)存在惟一解 f. 此时由基本定理可知这一个解：

$$f(x)=f(a)+\int_a^x g(s)\,\mathrm{d}s.$$

按 G. Strang(1991)，要证明右边积分存在，只要如3.5节那样，求它的上下和(要用连续函数的性质). 另一种证明则取于柯朗(2001)，证明欧拉算法(5-5)得到的序列 $\{f_n\}$ 在闭区间上是一致基本列：

$$\underset{[a,b]}{\text{upper}}|f_n(x)-f_m(x)|$$
$$\leqslant(b-a)\left[\epsilon_1\left(\frac{b-a}{n}\right)+\epsilon_1\left(\frac{b-a}{m}\right)\right].$$

先承认实数的一个性质(另册证明)：基本列必是收敛列.

既然 f_n、f_m 都是连续函数，那么由上述得到的一致收敛性可知其极限 f 仍是连续函数，并且由欧拉算法的构造可知 f 满足微分方程，从而基本定理成立.

设 $a=x_0<x_1<\cdots<x_n<x_{n+1}=b$，$h=x_{i+1}-x_i$，欧拉算法如下：

$$f_n(x)=f(a)+\sum_{0\leqslant s<i}g(x_s)h+g(x_i)(x-x_i),\quad x\in[x_i,x_{i+1}), \tag{3}$$

这样，针对该剖分，还得到一对应的分段常数函数

$$g_n(x)=g(x_{i-1}),\quad x\in[x_{i-1},x_i).$$

序列$\{f_n\}$是闭区间上的一致基本列的证明.

首先考虑以下两个剖分，剖分(a)是将区间$[a,b]$剖分成n个小段，剖分(b)是将区间$[a,b]$剖分成m个小段，然后将两个剖分的分点合并，得到一个新的剖分，记作剖分(c)，显然剖分(c)既是剖分(a)的加密，也是剖分(b)的加密. 利用这三剖分，我们容易构造各自的欧拉序列和对应的分段常数函数(图5-4).



图5-4 三个剖分

现在，换个角度来看剖分(c)对应的分段常数函数g_{m+n},利用剖分(c)是剖分(a)的加密这一性质,它在剖分(a)的每个小段上仍然是分段常数函数，并可以写成等价形式$\overline{g}$:

当x落在剖分(a)的第$i+1$个小段$[x_i,x_{i+1})$(其中还含有剖分(c)的节点$x_i=x_{i_1}<x_{i_2}<x_{i_3}<\cdots<x_{i_l}=x_{i+1}$),

$$\overline{g}(x)=\begin{cases}g(x_{i_1}), & x\in[x_{i_1},x_{i_2}),\\ g(x_{i_2}), & x\in[x_{i_2},x_{i_3}),\\ \quad\cdots\cdots & \\ g(x_{i_{l-1}}), & x\in[x_{i_{l-1}},x_{i_l}),\end{cases}$$

于是剖分(a)得到的欧拉序列f_n和剖分(c)得到的欧拉序列

f_{m+n}的差别是

$$f_n(x)-f_{m+n}(x)=\sum_{n\leqslant s\leqslant i-1}^{(a)}(g(x_s)-\bar{g})h_s+(g(x_i)-\bar{g})(x-x_i),$$

$x\in[x_i,x_{i+1})$.

其中，上标(a)表示针对剖分(a)求和，便有

$$\begin{aligned}|f_n(x)-f_{m+n}(x)|&\leqslant\sum_{0\leqslant s\leqslant i-1}^{(a)}|g(x_s)-\bar{g}|h_s+|g(x_i)-\bar{g}|(x-x_i)\\&\leqslant\max_{0\leqslant s\leqslant i}\underset{[x_s,x_{s+1}]}{\text{upper}}|g(\xi)-g(\eta)|(x-a),\quad x\in[x_i,x_{i+1}).\end{aligned}$$

从而

$$\begin{aligned}\underset{[a,b]}{\text{upper}}|f_n(x)-f_{m+n}(x)|&\leqslant(b-a)\max_{0\leqslant s\leqslant n}\underset{[x_s,x_{s+1}]}{\text{upper}}|g(\xi)-g(\eta)|\\&\leqslant(b-a)\epsilon_1\left(\frac{b-a}{n}\right).\end{aligned}$$

其中利用了 g 的一致连续性 (2).

同理，还可以得到

$$\underset{[a,b]}{\text{upper}}|f_m(x)-f_{m+n}(x)|\leqslant(b-a)\epsilon_1\left(\frac{b-a}{m}\right).$$

于是就有

$$\begin{aligned}&\underset{[a,b]}{\text{upper}}|f_n(x)-f_m(x)|\\&\leqslant\underset{[a,b]}{\text{upper}}|f_n(x)-f_{m+n}(x)|+\underset{[a,b]}{\text{upper}}|f_m(x)-f_{m+n}(x)|\\&\leqslant(b-a)\left[\epsilon_1\left(\frac{b-a}{n}\right)+\epsilon_1\left(\frac{b-a}{m}\right)\right].\end{aligned}$$

所以，当 n，m 充分大时，便得到序列 $\{f_n\}$ 是一致基本列. 于是就能找到一个函数 $f(x)$ 使得

$$\underset{[a,b]}{\text{upper}}|f(x)-f_n(x)|\leqslant\epsilon_2\left(\frac{b-a}{n}\right).\tag{4}$$

现在只剩下证明 f 满足微分方程(1). 既然对所有的 n, $f_n(a)=\alpha$, 显然 $f(a)=\alpha$. 因此，只需证明 $f'(x)=g(x)$, $x\in[a,b]$.

先证明 f_n 近似满足微分方程. 设 $a\leqslant x\leqslant b$, $a\leqslant x-h\leqslant b$, 并假设 $x_{i-1}<x-h\leqslant x_i$, $x_{j-1}\leqslant x\leqslant x_j$, 见图 5-5. 这里省略指标 n 使得记号简洁.

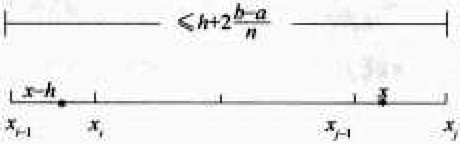


图 5-5 小区间

利用 f_n 的定义(3)和一致连续性(2)可知:

$$
\begin{aligned}
&|f_n(x)-f_n(x-h)-g(x)h| \\
&= \left|\sum_{s=i+1}^{j-1} g(x_{s-1})(x_s-x_{s-1})+g(x_{i-1})(x_i-x-h)\right. \\
&\quad \left. + g(x_{j-1})(x-x_{j-1})-g(x)h\right| \\
&= \left|\sum_{s=i+1}^{j-1}(g(x_{s-1})-g(x))(x_s-x_{s-1})\right. \\
&\quad + (g(x_{i-1})-g(x))(x_i-x-h) \\
&\quad \left. + (g(x_{j-1})-g(x))(x-x_{j-1})\right| \\
&\leqslant h\max_{[x_{i-1},x_j]}|g(z)-g(y)| \\
&\leqslant h\epsilon_1\left(h+2\frac{b-a}{n}\right). \qquad (5)
\end{aligned}
$$

由(4)和(5)便得到

$$
\begin{aligned}
&|f(x)-f(x-h)-g(x)h| \\
&\leqslant |f_n(x)-f_n(x-h)-g(x)h| + |f(x)-f_n(x)| \\
&\quad + |f(x-h)-f_n(x-h)| \\
&\leqslant h\epsilon_1\left(h+2\frac{b-a}{n}\right)+2\epsilon_2\left(\frac{b-a}{n}\right) \\
&\leqslant h\epsilon_3(h). \qquad (6)
\end{aligned}
$$

这里令 n 充分大，使得 $2\epsilon_2((b-a)/n)\leqslant h\epsilon_1(h)$，$2((b-a)/n)\leqslant h$.

由(6)可知 f 一致可导，且有 $f'=g$，这样，便找到了微分方程(1)的解．根据基本定理又可以得到

$$f(x)=\int_a^x g(s)\,\mathrm{d}s+f(a).$$

以上是由我的学生谢和虎和陈竑焘帮助推算的.

最后证明解惟一．假设有另外一个解 k，因为 $(f-k)'=g-g\equiv 0\Rightarrow f-k\equiv C$，并且 $f(a)=k(a)$，所以 $f\equiv k$，这就证明了惟一性.

这样便回答了本章一开头所提出的问题．我们把它们整理成下面的定理.

定理　假设 g 在 $[a,b]$ 上一致连续，那么存在惟一的可微函数

$$f(x)=\int_a^x g(s)\,\mathrm{d}s+\alpha$$

也称 g 的原函数，使 $f'=g$，$f(a)=\alpha$.

所以连续函数的积分存在，特别1.6节或3.8节的弧长存在.

第6章 多元微积分(并行推广)

世界是多元的.

多元微积分比一元复杂得多. 本章讲述可以照搬的那些部分.

6.1 格林公式

考虑定义在边界为 ∂D^+ 的矩形区域 $D=[a,b]\times[c,d]$ 上的二元函数（图6-1）：

格林公式 按导数定义（2-1），f 和 g 分别是 y 和 x 的偏可微的函数，于是

$$\oint_{\partial D^+} f(x,y)\,\mathrm{d}x+g(x,y)\,\mathrm{d}y=\iint_D\left(\frac{\partial g}{\partial x}-\frac{\partial f}{\partial y}\right)\mathrm{d}x\mathrm{d}y.$$

证明 这个二元的基本公式可以简化为一元的基本公式：

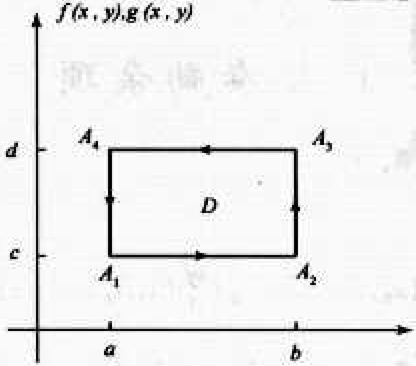


图6-1 有向图

$$\begin{aligned}\iint_D \frac{\partial g}{\partial x}\,\mathrm{d}x\mathrm{d}y &= \int_c^d\left(\int_a^b \frac{\partial g}{\partial x}\,\mathrm{d}x\right)\mathrm{d}y \\ &= \int_c^d [g(b,y)-g(a,y)]\,\mathrm{d}y \\ &= \int_c^d g(b,y)\,\mathrm{d}y - \int_c^d g(a,y)\,\mathrm{d}y \\ &= \int_{\overline{A_2A_3}} g(x,y)\,\mathrm{d}y + \int_{\overline{A_4A_1}} g(x,y)\,\mathrm{d}y \\ &= \int_{\overline{A_1A_2}+\overline{A_2A_3}+\overline{A_3A_4}+\overline{A_4A_1}} g(x,y)\,\mathrm{d}y \\ &= \oint_{\partial D^+} g(x,y)\,\mathrm{d}y\ ,\end{aligned}$$

即

$$\oint_{\partial D^+} g\,\mathrm{d}y = \iint_D \frac{\partial g}{\partial x}\mathrm{d}x\mathrm{d}y\,.$$

同理，

$$\oint_{\partial D^+} f\,\mathrm{d}x = -\iint_D \frac{\partial f}{\partial y}\mathrm{d}x\mathrm{d}y.$$

6.2 泰勒余项

考虑泰勒余项，如

$$
\begin{aligned}
& R(x,y) \\
& := f(x,y)-f(x_0,y_0)-(x-x_0)\frac{\partial f}{\partial s_1}(x_0,y_0)-(y-y_0)\frac{\partial f}{\partial t_1}(x_0,y_0) \\
& = f(x,y)-f(x,y_0)+f(x,y_0)-f(x_0,y_0) \\
& \quad -(x_0-x)\frac{\partial f}{\partial s_1}(x_0,y_0)-(y-y_0)\frac{\partial f}{\partial t_1}(x_0,y_0) \\
& = \int_{y_0}^{y}\frac{\partial f}{\partial t_1}(x,t_1)\,\mathrm{d}t_1+\int_{x_0}^{x}\frac{\partial f}{\partial s_1}(s_1,y_0)\,\mathrm{d}s_1-\int_{x_0}^{x}\frac{\partial f}{\partial s_1}(x_0,y_0)\,\mathrm{d}s_1 \\
& \quad -\int_{y_0}^{y}\frac{\partial f}{\partial t_1}(x_0,y_0)\,\mathrm{d}t_1 \quad \text{（其中用到图 6-2、图 6-3）} \\
& = \int_{y_0}^{y}\left[\frac{\partial f}{\partial t_1}(x,t_1)-\frac{\partial f}{\partial t_1}(x_0,y_0)\right]\mathrm{d}t_1+\int_{x_0}^{x}\int_{x_0}^{s_1}\frac{\partial^2 f}{\partial s_2^2}(s_2,y_0)\,\mathrm{d}s_2\mathrm{d}s_1 \\
& = \int_{y_0}^{y}\left[\frac{\partial f}{\partial t_1}(x,t_1)-\frac{\partial f}{\partial t_1}(x_0,t_1)\right]\mathrm{d}t_1 \\
& \quad +\int_{y_0}^{y}\left[\frac{\partial f}{\partial t_1}(x_0,t_1)-\frac{\partial f}{\partial t_1}(x_0,y_0)\right]\mathrm{d}t_1 \\
& \quad +\int_{x_0}^{x}\int_{x_0}^{s_1}\frac{\partial^2 f}{\partial s_2^2}(s_2,y_0)\,\mathrm{d}s_2\mathrm{d}s_1 \\
& = \int_{y_0}^{y}\int_{x_0}^{x}\frac{\partial^2 f}{\partial s_1\partial t_1}(s_1,t_1)\,\mathrm{d}s_1\mathrm{d}t_1+\int_{y_0}^{y}\int_{y_0}^{t_1}\frac{\partial^2 f}{\partial t_2^2}(x_0,t_2)\,\mathrm{d}t_2\mathrm{d}t_1 \\
& \quad +\int_{x_0}^{x}\int_{x_0}^{s_1}\frac{\partial^2 f}{\partial s_2^2}(s_2,y_0)\,\mathrm{d}s_2\mathrm{d}s_1 \quad \text{（其中用到图 6-4、图 6-5），}
\end{aligned}
$$

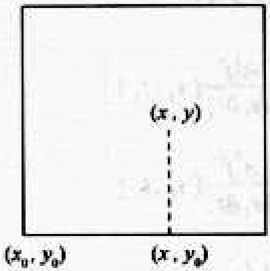


图 6-2　各点关系

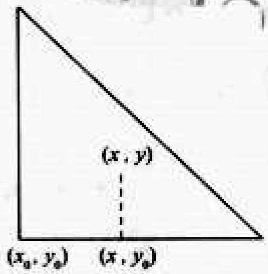


图 6-3　各点关系

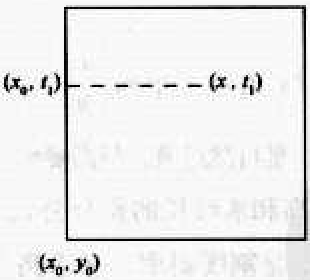


图 6-4　各点关系

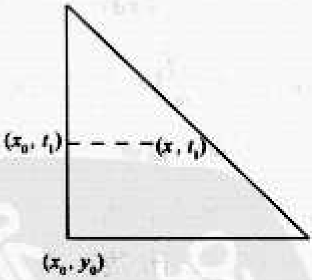


图 6-5　各点关系

$$
\begin{aligned}
|R| \leqslant & \frac{(x-x_0)^2}{2} \text{upper} \left| \frac{\partial^2 f}{\partial s_2^2}(s_2, y_0) \right| \\
& + (x-x_0)(y-y_0) \text{upper} \left| \frac{\partial^2 f}{\partial s_1 \partial t_1}(s_1, t_1) \right| \\
& + \frac{(y-y_0)^2}{2} \text{upper} \left| \frac{\partial^2 f}{\partial t_2^2}(x_0, t_2) \right| ,
\end{aligned}
$$

$$\left|\frac{\partial R}{\partial x}\right|+\left|\frac{\partial R}{\partial y}\right|\leqslant(x-x_0)\,\mathrm{upper}\left|\frac{\partial^2 f}{\partial s_2^2}(s_2,y_0)\right|$$
$$+(y-y_0)\,\mathrm{upper}\left|\frac{\partial^2 f}{\partial s_1\partial t_1}(s_1,t_1)\right|$$
$$+(x-x_0)\,\mathrm{upper}\left|\frac{\partial^2 f}{\partial s_1\partial t_1}(s_1,t_1)\right|$$
$$+(y-y_0)\,\mathrm{upper}\left|\frac{\partial^2 f}{\partial t_2^2}(x_0,t_2)\right|,$$

$$\left|\frac{\partial^2 R}{\partial x^2}\right|+\left|\frac{\partial^2 R}{\partial x\partial y}\right|+\left|\frac{\partial^2 R}{\partial y^2}\right|$$

$$\leqslant\mathrm{upper}\left|\frac{\partial^2 f}{\partial s_2^2}(s_2,y_0)\right|+\mathrm{upper}\left|\frac{\partial^2 f}{\partial s_1\partial t_1}(s_1,t_1)\right|+\mathrm{upper}\left|\frac{\partial^2 f}{\partial t_2^2}(x_0,t_2)\right|.$$

以上证明有点繁，但中间并无跨度，是自然过渡，只需要耐心.

在一元情形，我们已经介绍了求高和求弧长的积分公式. 在那里，我们利用切线斜率 $f'(x)$，而不是割线斜率，来求高 $f(x)$ 及其弧长. 现在看看如何将这个切线方法应用到多元微积分上，即如何用切线斜率来求二元情形的高 $f(x,y)$，以及曲面面积.

6.3 求高公式

如果曲面表达成函数 $f(x,y)$（定义在区域 $[a,x]\times[b,y]$ 上），那么曲面在 P、Q 点的高度差可表达成（图6-6）

$$f(Q)-f(P)=f(R)-f(P)+f(Q)-f(R),$$

即横向和竖向高度差的和. 这时只需要分别使用(3-8)式，即可得到

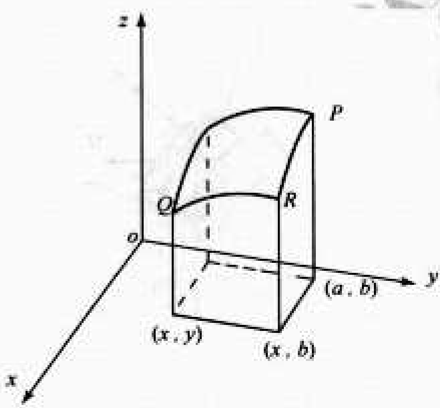


图6-6　曲面高度差

$$f(Q)-f(P)=\int_a^x \frac{\partial f}{\partial x}(x,b)\,\mathrm{d}x+\int_b^y \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}y.$$

这便是与一元函数相类似的“二元变上限积分”公式.

6.4　求面积公式

再次采用微元法，把曲面分片，在一片上，曲面被换成切平面（图6-7），即

$$S_{ABCD}:=S_{AMPN},$$

而不是割平面. 上面的几何图形过于复杂，为了看清楚它的本来面目，我们把横向和竖向的小曲面及小切平面放大，并加上相应坐标.

现在我们再做一个简化即 $AMPN$ 为一个平行四边形. 接下来就是在三维空间中计算 $AMPN$ 的面积了. 利用中学的求面积公式，可以计算出

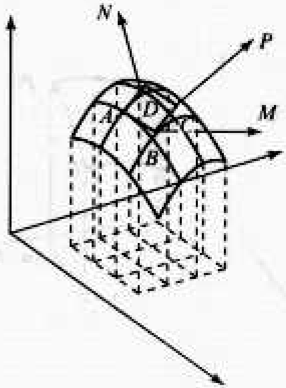


图 6-7 曲面与切平面

$$S_{AMPN} = \sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}\Delta x \Delta y,$$

下面只需要再把它们加起来，就有[①]

$$S = \iint_D \sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}\,\mathrm{d}x\mathrm{d}y.$$

多元微积分，比起一元情形，难在缺少几何直观.

① 这是我的学生周俊明帮助算的.

第 7 章 轻松的抽象微积分(照猫画虎)

先有猫再有虎.

照猫画虎：泛函空间≈平面几何. 这也就是为什么数学家关肇直说“泛函空间是平面几何”. 我们可以一眼认出泛函分析的真面目.

7.1　函数和向量

函数空间（赋范空间和内积空间）对于初学者来讲过于抽象，如何认出它的真面目呢？其实，它在平面几何，例如两点以直线为最短，还有三角形的余弦定理中，就已经见过. 所以，可能通过平面几何来认出函数空间，这里以几何的思想代替分析的技巧，一眼看穿若干分析不等式的必然性，多轻松！

7.1.1 赋范空间的发明

中学已开始学习空间向量. 那里，一个向量由三个分量刻画：

$$f=(f_1, f_2, f_3),$$

并带有长度. 依勾股定理

$$\|f\|=\sqrt{\sum_{i=1}^{3} f_i^2}.$$

第二个向量

$$g=(g_1, g_2, g_3),$$

一起形成了一个三角形，见图 7-1.

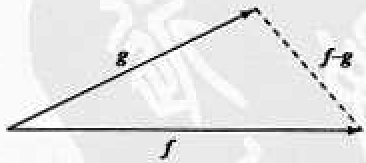


图 7-1 两条向量

其中第三边为向量

$$f-g=(f_1-g_1, f_2-g_2, f_3-g_3).$$

依第三边小于两边之和（两点间以直线段为最短），有三角不等式

$$\|f-g\|\leqslant\|f\|+\|g\|.$$

写成代数形式为

$$\sqrt{\sum_{i=1}^{3}(f_i-g_i)^2}\leqslant\sqrt{\sum_{i=1}^{3}f_i^2}+\sqrt{\sum_{i=1}^{3}g_i^2}$$

(仅当三点一线时取等号).

代数语言启发我们发明多个分量的“向量”(这里保留几何名称),以及建立它们之间的关系. 例如,当

$$\boldsymbol{f}=(f_1,\ f_2,\ \cdots,\ f_n),\quad \boldsymbol{g}=(g_1,\ g_2,\ \cdots,\ g_n),$$

作为类比,这两个多分量的向量也应该各有长度

$$\|\boldsymbol{f}\|=\sqrt{\sum_{i=1}^{n}f_i^2},\quad \|\boldsymbol{g}\|=\sqrt{\sum_{i=1}^{n}g_i^2}.$$

根据第三边小于两边之和,感觉也有三角不等式

$$\sqrt{\sum_{i=1}^{n}(f_i-g_i)^2}\leqslant\sqrt{\sum_{i=1}^{n}f_i^2}+\sqrt{\sum_{i=1}^{n}g_i^2}.$$

这是武断的发明,以后再给予严格、规范的证明. 感觉在前,严格在后.

转折点在于把函数也视为向量. 先看一个 n 段的折线函数(图 7-2).

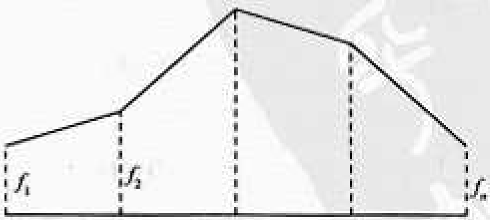


图 7-2　折线函数

它由 n 个分量刻画:

$$\boldsymbol{f}=(f_1,\ f_2,\ \cdots,\ f_n).$$

让分量越来越多，极端状态便视为一个函数（图 7-3）.

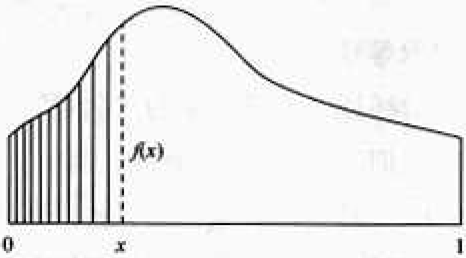


图 7-3　一般函数

它由无数个分量刻画：

$$\boldsymbol{f}=(f(t),\quad 0\leqslant t\leqslant 1),$$

即函数也可以视为一个向量，只是它有无数个分量 $f(t)$，t 从 0 跑遍 1.

一旦函数 f 被视为向量，其长度类比为

$$\|\boldsymbol{f}\|=\sqrt{\int_0^1 f(t)^2\mathrm{d}t}.$$

另一函数 g 也是向量，

$$\boldsymbol{g}=(g(t),\quad 0\leqslant t\leqslant 1),$$

第三个向量

$$\boldsymbol{f}-\boldsymbol{g}=(f(t)-g(t),\quad 0\leqslant t\leqslant 1)$$

的长度也应该小于前两条的长度之和，三角不等式变成了闵可夫斯基不等式：

$$\sqrt{\int_0^1[f(t)-g(t)]^2\mathrm{d}t}\leqslant\sqrt{\int_0^1 f(t)^2\mathrm{d}t}+\sqrt{\int_0^1 g(t)^2\mathrm{d}t}.$$

这也是先一眼认出，再给予严格、规范的证明. 感觉在前，严格

在后.

所以，大学(函数、积分)和中学(向量、求和)有何等的默契!

先具体后抽象：赋范空间就是每个向量有长度(又称范数)，两个向量之间有三角不等式，后者反映了两点以直线为最短的几何性质. 详见 7.1.3 小节最后.

7.1.2　内积空间的发明

两个向量，f 和 g，有夹角 θ（图 7-4）.

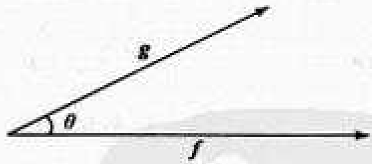


图 7-4　夹角

由余弦定理得

$$\|f-g\|^2=\|f\|^2+\|g\|^2-2\|f\|\|g\|\cos\theta,$$

夹角余弦为

$$\cos\theta=\frac{\|f\|^2+\|g\|^2-\|f-g\|^2}{2\|f\|\|g\|}.$$

写成代数：

$$\cos\theta=\frac{\sum f_i g_i}{\sqrt{\sum f_i^2}\sqrt{\sum g_i^2}}$$

或

$$\cos\theta = \frac{\int_0^1 f(t)g(t)\,\mathrm{d}t}{\sqrt{\int_0^1 f(t)^2\,\mathrm{d}t}\sqrt{\int_0^1 g(t)^2\,\mathrm{d}t}}.$$

分子为新出现的数量，成为向量$\boldsymbol{f}$和$\boldsymbol{g}$的内积. 假如在多个分量甚至无数个分量的空间之中，仍然接受

$$|\cos\theta| \leqslant 1.$$

我们就必须接受施瓦茨不等式

$$\left|\sum f_i g_i\right| \leqslant \sqrt{\sum f_i^2}\sqrt{\sum g_i^2}$$

或

$$\left|\int_0^1 f(t)g(t)\,\mathrm{d}t\right| \leqslant \sqrt{\int_0^1 f(t)^2\,\mathrm{d}t}\sqrt{\int_0^1 g(t)^2\,\mathrm{d}t}.$$

这也是先一眼认出，再给予严格、规范的证明. 感觉在前，严格在后.

所以，大学(函数、积分)和中学(向量、求和)有何等的默契！

内积空间就是两个向量有内积(详见7.1.3小节)，它反映了角度的余弦≤1(指绝对值).

7.1.3 证 明

为了证明施瓦茨(和闵可夫斯基)不等式，引入一个简便的概念：

$$(\boldsymbol{f},\boldsymbol{g}) = \sum f_i g_i = \int_0^1 f(t)g(t)\,\mathrm{d}t$$

称为$\boldsymbol{f}$和$\boldsymbol{g}$的（实的）内积，它类似于实数的乘积：

$$(f,f)\geqslant 0,\qquad (f,f)=0\Leftrightarrow f=0,$$

$$(f,g)=(g,f),$$

$$(f+w,g)=(f,g)+(w,g),$$

$$(\alpha f,g)=\alpha(f,g),\qquad \alpha\text{为任意实数}.$$

这些“实数乘积”就能保证施瓦茨(和闵可夫斯基)不等式. 为此，首先定义一个由内积诱导出的长度，又称范数 $\|\cdot\|$：

$$(f,f)=\|f\|^2.$$

然后，对于单位向量 f 和 g：

$$\|f\|=\|g\|=1.$$

根据内积的性质：

$$\|f\pm g\|^2=\|f\|^2\pm 2(f,g)+\|g\|^2=2[1\pm(f,g)]\geqslant 0,$$

有

$$|(f,g)|\leqslant 1.$$

对于非单位的向量 f 和 g，根据内积的最后一个性质，也能够规格化：

$$\left|\left(\frac{f}{\|f\|},\frac{g}{\|g\|}\right)\right|\leqslant 1$$

或者

$$|(f,g)|\leqslant\|f\|\|g\|.$$

这就证明了施瓦茨不等式. 这里以代数的演算代替分析的技巧，只有7行. 长了记不住!

施瓦茨不等式又导出闵可夫斯基不等式，或三角不等式

$$\|f-g\|^2=\|f\|^2-2(f,g)+\|g\|^2$$

$$\leqslant\|f\|^2+2\|f\|\|g\|+\|g\|^2$$

$$=(\|f\|+\|g\|)^2,$$

所以，内积空间诱导出赋范空间，长度或范数满足

$$\|f\|\geqslant 0,\quad \|f\|=0\Leftrightarrow f=0,$$

$$\|\alpha f\|=|\alpha|\,\|f\|,\quad \forall\alpha\in\mathbf{R},$$

$$\|f+g\|\leqslant\|f\|+\|g\|.$$

7.1.4 正 交 性

一个重要特例就是两个向量 f 和 g 正交：

$$(f,g)=0\Leftrightarrow\cos\theta=0.$$

这时有勾股公式（图 7-5）：

$$\|f-g\|^2=\|f\|^2+\|g\|^2.$$

没有正交条件，只能够得到平行四边形恒等式：

$$\|f-g\|^2=2\|f\|^2+2\|g\|^2-\|f+g\|^2$$

这对于泛函分析中某些收敛性（或存在性）的精确论证是关键（图 7-6）.

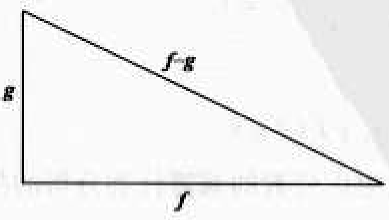


图 7-5　直角三角形

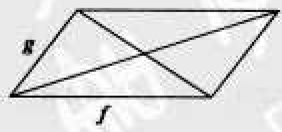


图 7-6　平行四边形

正交的概念很有用．称 $\varphi\in V$ 是向量 f 在子空间 V 上的正交投影，如果满足：

$$(f-\varphi,g)=0,\quad\forall g\in V,$$

那么，感觉 φ 应该是子空间 V 中最靠近 f 的向量（图7-7）：

$$\|\varphi-f\|\leqslant\|g-f\|,\quad \forall g\in V.$$

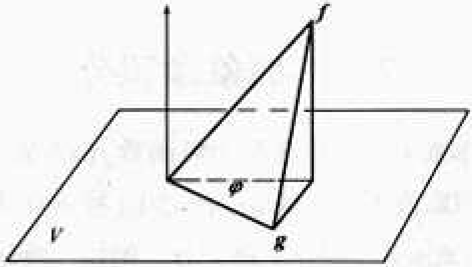


图7-7　正交和最小性一眼看出

证明只要两行：基于正交性

$$(f-\varphi,\ g-\varphi)=0,$$

变得 $\forall g\in V$，

$$\|f-g\|^2=\|f-\varphi\|^2+\|g-\varphi\|^2\geqslant\|f-\varphi\|^2.$$

上述正交和最小性质使偏微分方程的有限元算法变成了几何学（其中 f 是微分方程的真解，V 是有限元空间，φ 是有限元的解，离真解最近，g 也可以是有限元插值）. 所以，学习一点泛函分析的思想，有助于对分析学的理解，值得！

到此，都是照猫(平面几何)画虎(泛函空间)，所以有：

泛函空间≈平面几何.

但是完整的泛函分析和平面几何的本质区别在于其完备性：

泛函空间－平面几何＝完备性.

（它使更多的微分方程获得“广义解”）. 但完备性难，幸亏的是本章主题(见7.2节)不用它！

也可以参看文献亚历山大洛夫等(2001)著的《数学——它的内容,方法和意义》的第十九章.

7.2 抽象微积分

现在回到本章主题. 什么是抽象函数$f(x)$呢? 它的自变量x仍然是实数, 但取值于线性赋范空间之中(特例就是实变量的向量函数). 对于抽象函数, 仍然有微积分. 例如, 微分还是(2-8):

$$f(x+h)-f(x)=f'(\xi)h+\bar{\epsilon}(\xi,h)h,$$

其中相对误差按范数(代替绝对值)

$$\bar{\epsilon}(\xi,h)\ll 1.$$

积分或基本定理还是(3-6):

$$f(b)-f(a)-\sum_{\xi\in\text{中间点}} f'(\xi)h=(b-a)(\bar{\epsilon}\text{的平均}),$$

$$\bar{\epsilon}(\xi_i,h_i)\text{的平均}\leqslant\epsilon(h)\ll 1$$

(按范数).

注 一些泛函分析教材(如 Ljusternik(1965))含有抽象微积分的基本定理, 其证明基于 Hahn-Banach 定理和赋范空间的完备性, 然后还要利用实函数的微积分定理, 所以完整的证明是非常长的. 本书避开这一切难点, 直来直去, 几行推出抽象函数的基本定理, 多轻松!

附录 微分方程计算四步曲

"数学太难了", 数学家华罗庚说. 德国人提出专业演讲的

四步曲：第一步中学生能懂，第二步大学生懂，第三步研究生和专家才懂（那是专业），第四步只有你自己懂（那是最新成果），或连你自己也不懂（那是未来发展）．下面以微分方程的计算为例．

一、中学三角方程

已知斜率

$$\frac{f(x)}{x}=\tan\theta=g,$$

求树高：$f(x)=g\cdot x$，其中 x 是一个数（图 7-8）．

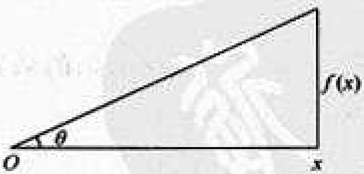


图 7-8　直角三角形

这个三角方程何用呢？不必砍树也能求树高！

二、大学微分方程

有三种方程．

1. 第一方程

已知山坡斜率（图 7-9）

$$f'(x)=g(x),\quad f(a)=\alpha. \tag{1}$$

求山高：$f(x)=?$

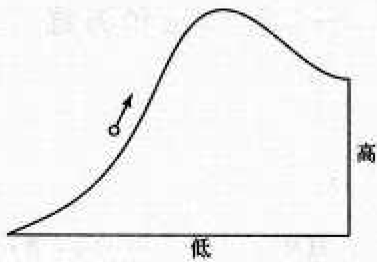


图 7-9 曲边三角形

其中，f 是函数，不是数．这就是最简单的微分方程，是中学三角的延伸，有着最明显的几何说明．

先把区间 $[a,x]$ 分为 $n+1$ 小段：

$$a=x_0<x_1<\cdots<x_n<x_{n+1}=x,\qquad h_i=x_{i+1}-x_i,$$

再把曲线小段用切线代替（包络），然后逐段求高（图 7-10）．

看图识字：

$$f_0=f(a),$$

$$f_1=f_0+g(x_0)h_0,$$

$$f_2=f_1+g(x_1)h_1,$$

$$f_{n+1}=f_n+g(x_n)h_n=f_0+\sum_{0\leqslant i<n+1}g(x_i)h_i$$

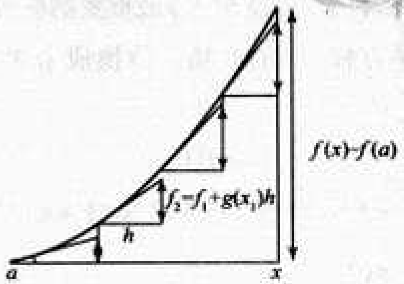


图 7-10　欧拉折线求高法一眼看出

称为欧拉折线法，按(3-4)有误差

$$f(x)-f_{n+1}=f(x)-f(a)-\sum_{0\leqslant i<n+1}f'(x_i)h_i\approx Ch \qquad (2)$$

为一阶精度，对微积分已足够：

$$f(x)-f(a)=\int_a^x f'(s)\,\mathrm{d}s .$$

对积分计算：令 $f'=g$，(2)有更确切的估计：

$$\int_a^x g(s)\,\mathrm{d}s-\sum_{0\leqslant i\leqslant n}g(x_i)h_i\leqslant\frac{1}{2}\sum_{0\leqslant i\leqslant n}h_i^2\ \underset{[x_i,x_{i+1}]}{\text{upper}}|g'|$$

(按绝对值)，由此导致最优网格(当 g' 大，取 h_i 小)．但是一阶精度还必须改进，如中点(或梯形)公式，能有二阶精度；辛普森公式(中点和梯形公式的线性组合)，能有四阶精度(见第 4 章附录)．

2. 第二方程

变化依赖于现状：

$$f'(x)=f(x)\ (x>0),\quad f(0)=1.$$

这是应用数学(生命科学,经济学等)最重要的微分方程. 也用欧拉折线法. 因为第一方程(1)的右端 g 被换成未知的 f, 所以 $g(x_n)$ 只能换成已知的 f_n:

$$f_0=f(0),$$

$$f_{n+1}=f_n+f_nh=(1+h)f_n=(1+h)^{n+1}f_0$$
$$=(1+h)^{\frac{x}{h}}$$

(用 $(n+1)h=x$),按(5-10)有误差

$$f(x)-f_{n+1}=e^x-(1+h)^{\frac{x}{h}}\leqslant Ch$$

(证明见(5-10))为一阶精度,对微积分已足够: $f'(x)=f(x)$(其中 $f(x)=e^x$ 是第二方程的解). 对数值计算, 精度必须改进, 见泰勒级数(4.4节).

何用? 用于人口普查: 一人几分钟普查代替了亿户一年普查(详见5.1.2小节)!

推广: 更一般的情形, 包括第一、第二方程, 为一阶常微分方程:

$$f'(x)=g(x,f),\quad f(a)=\alpha.$$

欧拉折线法其实已含在第二方程之中:

$$f_0=f(a),$$

$$f_{n+1}=f_n+g(x_n,f_n)h_n\to f(x)?$$

只可能有一阶精度, 不能预报天气, 必须改进, 如梯形公式, 能有二阶精度.

3. 第三方程

二阶方程

$$f''(x)=g(x),\quad f(a)=\alpha,\quad f(b)=\beta \tag{3}$$

出现在物理、工程(牛顿定律是加速度,不是速度)中．通过两端乘上试探函数，再分部积分，变成一阶形式，然后用折线法(又称有限元法)求解，恰好它是解的插值(在结点绝对精确)(图 7-11).

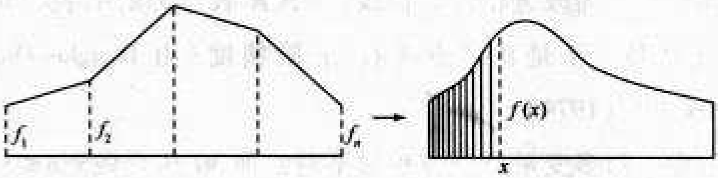


图 7-11

三、专业(多变量微分方程)

多变量复杂多了，只有有限元法是通用解法(图 7-12)．详见 7.1.4 节．但在结点不可能精确．精度必须改进，见下一段.

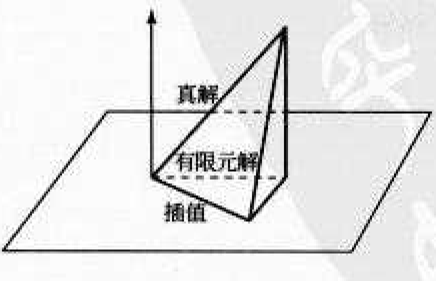


图 7-12　有限元法

有限元解是真解的正交投影，离真解最近（比插值更近）.

四、最新成果及未决问题

第三方程(3)的折线法在结点绝对精确，可惜对一般常微分方程不成立. 为改进精度，折线(一次函数)必须用高次(p 次)多项式代替，于是在结点能有 $2p$ 阶精度(由 Douglas-Dupont-Wheeler 发现(1974)).

问题：对多变量微分方程对不对? 前 30 年只做到$p\leq3$，最近才有人(周俊明,朱起定和魏继东等)向 $p=4$ 进军. 对于更大的 p，尚是悬案.

这个证明非常费劲，常人无法读懂. 有一种间接的高精度方法，是将有限元解跟某种插值相比较，它们之间的误差便有高精度. 这一方法在 C. Grossmann-H. G. Roos-M. Stynes(2007)的教材中已有证明，相当简单，有可读性. 也见 H. G. Roos-M. Stynes-L. Tobiska (2008) 的专著. 此外，更一般的讨论见 J. Brandts-M. Krizek(2001)的综述.

参考文献

阿诺德．2006．新蒙昧主义与俄罗斯教育．数学译林，（2）：108－121

布朗．1978．微分方程及其应用．纽约：施普林格

法国教育部．2001．法国数学课程标准．巴黎．Program of Higher Education Mathematics and Class Series of Scientifc. Done at Paris，66

华罗庚．1963．高等数学引论．北京：科学出版社

拉克斯等．1980．微积分及其应用与计算．北京：人民教育出版社

柯朗，约翰．2001．微积分和数学分析引论（第一卷第一分册）．北京：科学出版社

林群．1997．数学也能看图识字．光明日报，06-27；人民日报，08-06

林群．1999．画中漫游微积分．南宁：广西师范大学出版社

林群，吴从炘．2002．大学文科数学．保定：河北大学出版社

陶哲轩．2008．陶哲轩实分析．北京：人民邮电大学出版社

托尔斯泰．1987．战争与和平．北京：人民文学出版社

文丽．1981．一元函数积分学．上海：上海科技出版社

吴文俊．1995．论数学机械化．济南：山东教育出版社

亚历山大洛夫．2001．数学——它的内容，方法和意义．秦元勋等译．北京：科学出版社

於崇华等．1999．数学分析．北京：高等教育出版社

张景中．2008．微积分基础的新视角．中国科学，1：1～3

朱学贤，郑建华．2000．一元微积分．北京：高等教育出版社

卓里奇．2006．数学分析．蒋铎等译．北京：高等教育出版社

Arnold N. 1995. Graphics for the calculus classroom. [http://www. ima. umn. edu/~arnold/graphics. html]

Brandts J, Krizek M. 2001. History and future of superconvergence. JAKUTO International series, Math. Sci. Appl., vol. 15

Brunner H. 2008. Private communication

Bruter C. 1973. Sur La Nature Des Mathematiques. Paris: Gauthier-Villars

Dovermann K. 1999. Applied Calculus. [http://www. math. hawaii. edu/7%Eheiner/ calculus. pdf]

Grossmann C, Roos H G, Stynes M. 2007. Numerical Treatment of Partial Differential Equations. Springer-Verlag Berlin Heideberg

Karcher H. 2002. Analysis mit gleichmgen Fehlerschranken. Okt [http://www. math. uni-bonn. de/people/karcher]

Leger A. 2009. Pivate communication

Lin Q. 2008. Free Calculus: A Liberation from Concepts and Proofs. Singapore: World Scientific Pub

Livshits M. 2004. Simplifying Calculus by Using Uniform Estimates. [mysite. verizon. net/michaelliv www. mathfoolery. org]

Livshits M. 2008. Rethinking Calculus. [http://arxiv. org/abs/0905. 3611]

Ljusternik L, Sobolev V. 1965. Elements of Functional Analysis. Moscow

Papanicolaou G. 1973. Stochastic equations and their applications. AMM, 80 (5): 526~545

Roos H G, Stynes M, Tobiska L. 2008. Robust Numerical Methods for Singularly Perturbed Differential Equations. Springer-Verlag Berlin Heideberg

Ryan M. 2003. Calculus for Dummies. New York: Wiley Publishing Inc

Strang G. 1991. Calculus. Wellesley MA：Wellesley-Cambridge Press

Vainberg M. 1946. A Hamerstein theorem for nonlinear integral equations. Moscow Univ. Trans，100（1）：93－103

Zeldovich Y，Yaglom I. 1987. Higher Math for Beginners.［http：//rapids hare. com］